

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:53:30 ; Search time 12.56 Seconds

(without alignments)  
562,582 Million cell updates/sec

Title: US-09-912-717-1

Perfect score: 1567  
Sequence: 1 MSVFGAGQLAIFRTAAGL.....TPSSPGKLLTRSLALGCKKD 314

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :

1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCFUS.COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1567	100.0	314	US-09-099-676-1	Sequence 1, Appl1
2	1567	100.0	314	US-09-565-910-1	Sequence 1, Appl1
3	1318	84.1	315	US-09-099-676-3	Sequence 3, Appl1
4	1318	84.1	315	US-09-565-910-3	Sequence 3, Appl1
5	294	16.8	304	US-08-605-150A-2	Sequence 2, Appl1
6	102	6.5	349	US-08-605-150A-17	Sequence 17, Appl1
7	98	6.3	367	US-08-605-150A-6	Sequence 6, Appl1
8	98	6.3	381	US-08-375-109-17	Sequence 17, Appl1
9	97.5	6.2	543	US-08-752-929-17	Sequence 17, Appl1
10	97.5	6.2	543	US-09-090-793-10	Sequence 10, Appl1
11	97.5	6.2	543	US-08-726-012B-2	Sequence 2, Appl1
12	95.5	6.1	1455	US-08-605-150A-8	Sequence 8, Appl1
13	94.5	6.0	372	US-08-605-150A-10	Sequence 10, Appl1
14	94.5	6.0	372	US-08-614-377A-7	Sequence 7, Appl1
15	91.5	5.8	1026	US-09-142-648B-7	Sequence 7, Appl1
16	91.5	5.8	1026	US-08-961-083-84	Sequence 84, Appl1
17	88.5	5.6	722	US-08-605-150A-2	Sequence 2, Appl1
18	88.5	5.6	722	US-08-176-413-5	Sequence 5, Appl1
19	85	5.4	398	US-08-804-227C-14	Sequence 14, Appl1
20	85	5.4	398	US-08-804-227C-8	Sequence 8, Appl1
21	84	5.4	4550	US-08-804-227C-14	Sequence 14, Appl1
22	84	5.4	4550	US-08-804-198-2	Sequence 2, Appl1
23	83.5	5.3	801	US-08-725-012-2	Sequence 2, Appl1
24	83.5	5.3	801	US-08-264-534-34	Sequence 34, Appl1
25	83	5.3	657	US-08-083-350A-13	Sequence 13, Appl1
26	83	5.3	657	US-08-465-500-34	Sequence 34, Appl1
27	83	5.3	657	US-08-465-500-34	Sequence 34, Appl1

28	83	5.3	657	2	US-08-346-128-34	Sequence 34, Appl1
29	83	5.3	657	3	US-08-532-384-13	Sequence 13, Appl1
30	83	5.3	657	3	US-08-893-828-34	Sequence 34, Appl1
31	83	5.3	1015	1	US-08-537-210A-1	Sequence 1, Appl1
32	83	5.3	1015	1	US-09-113-825-16	Sequence 16, Appl1
33	83	5.3	2471	1	US-08-185-432-16	Sequence 16, Appl1
34	83	5.3	2471	1	US-08-083-590A-19	Sequence 19, Appl1
35	83	5.3	2471	3	US-08-532-384-19	Sequence 19, Appl1
36	82	5.2	1093	3	US-08-545-860D-55	Sequence 55, Appl1
37	82	5.2	1093	3	PCF-US94-04496-55	Sequence 55, Appl1
38	81.5	5.2	371	4	US-09-347-803-27	Sequence 27, Appl1
39	81.5	5.2	1026	1	US-08-194-290-7	Sequence 7, Appl1
40	80.5	5.1	530	1	US-08-187-793-4	Sequence 4, Appl1
41	80	5.1	417	4	US-09-355-115-2	Sequence 2, Appl1
42	79.5	5.1	427	4	US-08-886-634-2	Sequence 2, Appl1
43	79.5	5.1	427	4	US-09-349-851-2	Sequence 2, Appl1
44	79	5.0	425	3	US-08-946-329A-19	Sequence 19, Appl1
45	79	5.0	425	4	US-08-567-357A-19	Sequence 19, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-099-676-1  
Sequence 1, Application US/09099676  
Patent No. 6100075  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Corley, Neil C.  
APPLICANT: Baughn, Mariah R.  
TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESS: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/099,676  
FILING DATE: HERWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cerrone, Michael C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0532 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-855-0572  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 314 amino acids  
TYPE: amino acid  
STRANDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PROSNO1  
CLONE: 2278458  
US-09-099-676-1

Query Match 100.0%; Score 1567; DB 3; Length 314;  
Best Local Similarity 100.0%; Pred. No. 2.5e-163;  
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVGFAGOLAYRTAGILSAHKIIISPEMNLPTVSALKRMGVNLTFRSKETVKHSD 60  
DB 1 MSVGFAGOLAYRTAGILSAHKIIISPEMNLPTVSALKRMGVNLTFRSKETVKHSD 60  
QY 61 VLFIAVKPHIIPFLIDEIGADVQARHIIVSCAAGVTISSVEKKIMAFQAPAVIRCMNT 120  
DB 61 VLFIAVKPHIIPFLIDEIGADVQARHIIVSCAAGVTISSVEKKIMAFQAPAVIRCMNT 120  
QY 121 PVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGPAYAFMA 180  
DB 121 PVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGPAYAFMA 180  
QY 181 LDADGVKMGIPRRALIQGAQALLGAAKMLDSEQHPQKLDNVCSFGATIHAFLE 240  
DB 181 LDADGVKMGIPRRALIQGAQALLGAAKMLDSEQHPQKLDNVCSFGATIHAFLE 240  
QY 241 SGGFRSLINAVEASCIRTELQSMADQEKISPAALKKTLIDRYKLESPVSTLTPSSPG 300  
DB 241 SGGFRSLINAVEASCIRTELQSMADQEKISPAALKKTLIDRYKLESPVSTLTPSSPG 300  
QY 301 KLTSLALGKKD 314  
DB 301 KLTSLALGKKD 314

## RESULT 2

US-09-565-940-1  
Sequence 1, Application US/09565910  
Patent No. 6288192

## GENERAL INFORMATION

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Baughn, Mariah R.

TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/565, 910

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/099, 676

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Cerrone, Michael C.

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PF-0532 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-855-0572

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 314 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:  
LIBRARY: PROSNON01  
CLONE: 2278458  
US-09-565-910-1

Query Match 100.0%; Score 1567; DB 4; Length 314;  
Best Local Similarity 100.0%; Pred. No. 2.5e-163;  
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVGFAGOLAYRTAGILSAHKIIISPEMNLPTVSALKRMGVNLTFRSKETVKHSD 60  
DB 1 MSVGFAGOLAYRTAGILSAHKIIISPEMNLPTVSALKRMGVNLTFRSKETVKHSD 60  
QY 61 VLFIAVKPHIIPFLIDEIGADVQARHIIVSCAAGVTISSVEKKIMAFQAPAVIRCMNT 120  
DB 61 VLFIAVKPHIIPFLIDEIGADVQARHIIVSCAAGVTISSVEKKIMAFQAPAVIRCMNT 120  
QY 121 PVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGPAYAFMA 180  
DB 121 PVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGPAYAFMA 180  
QY 181 LDADGVKMGIPRRALIQGAQALLGAAKMLDSEQHPQKLDNVCSFGATIHAFLE 240  
DB 181 LDADGVKMGIPRRALIQGAQALLGAAKMLDSEQHPQKLDNVCSFGATIHAFLE 240  
QY 241 SGGFRSLINAVEASCIRTELQSMADQEKISPAALKKTLIDRYKLESPVSTLTPSSPG 300  
DB 241 SGGFRSLINAVEASCIRTELQSMADQEKISPAALKKTLIDRYKLESPVSTLTPSSPG 300  
QY 301 KLTSLALGKKD 314  
DB 301 KLTSLALGKKD 314

## RESULT 3

US-09-099-676-3  
Sequence 3, Application US/09099676  
Patent No. 6100075

## GENERAL INFORMATION

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Baughn, Mariah R.

TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/099, 676

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Cerrone, Michael C.

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PF-0532 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-855-0572

TELEX:

```

? INFORMATION FOR SEQ ID NO: 3 :
? SEQUENCE CHARACTERISTICS:
? LENGTH: 315 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? IMMEDIATE SOURCE:
? LIBRARY: GENBANK
? CLONE: 189498
?
US-09-099-676-3

```

Query Match	84.1%;	Score 1318;	DB 3;	Length 315;
Best Local Similarity	83.8%;	Pred. No. 4.7e-136;		
Matches 263; Conservative	27;	Mismatches 22;	Indels 2;	Gaps 1;

Qy	1	MSVGFAGQOLVYR--ETAAAGTSLSHNHKTIASSPENKILPTVYSALEKMGVNTJRNKEETVH	58
Db	1	MSVGFAGQOLFAKGFMAVGLAANHIMASSPDMDLATVYSALEKMGVKLTLPINKETVQH	60
Qy	59	SDVLFIAVVRPHIIPITLIDEGADVOARITVYSCAAGTIVSSVEKKLMAFOPARVYTRCMT	118
Db	61	SDVLFIAVVRPHIIPITLIDEGADIDIRHIVYSCAAGTIVSSIEKKLSAFRPARVYTRCMT	120
Qy	119	NTPVYVVOEGATVYATGTTHALVYEDGQLLEQLMSVGFCTEVEEDLIDAVTGLSSGPAVAF	178
Db	121	NTPVYVVRGEGATVYATGTTHAQVYEDGRMLGQLLTSTVGFCTVEEEDLIDAVTGLSSGPAVAF	180
Qy	179	MALDADGCGVKKMLPRRLAIQLGSAQLLGAAMLLDSEQHPCQLKDNVCSFGGATTHALHF	238
Db	181	TALDADGCGVKKMLPRRLAVRLGAQQLLGAAMLLHSHQHPCQLKDNVCSFGGATTHALHY	240
Qy	239	LESSEGFRLILINAVENASCIPTRELOSMADQEKISPAALKTLLDRVYLESPVYSTLTLPSS	298
Db	241	LESSEGFRLILINAVENASCIPTRELOSMADQGVYVPAIKKTIIDKYLDSPATALSPSG	300
Qy	299	PGKLLTRSLALGK 312	
Db	301	HTKLLPRSLAPAGK 314	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

:
:
:   REGISTRATION NUMBER: 39,132
:   REFERENCE/DOCKET NUMBER: PF-0532 US
:
:   TELECOMMUNICATION INFORMATION:
:
:   TELEPHONE: 650-855-0555
:   TELEFAX: 650-855-0572
:
:   TELEX:
:
:   INFORMATION FOR SEQ ID NO: 3:
:
:   SEQUENCE CHARACTERISTICS:
:
:     LENGTH: 315 amino acids
:     TYPE: amino acid
:     STRANDEDNESS: single
:     TOPOLOGY: linear
:
:   IMMEDIATE SOURCE:
:   LIBRARY: GENBANK
:   CLONE: 189498
:
:
: US-09-565-910--3

```

Query Match	84.1%;	Score 1318;	DB 4;	Length 315;
Best Local Similarity	83.8%;	Pred. No. 4.7e-136;		
Matches 263;	Conservative 27;	Mismatches 22;	Indels 2;	Gaps 1;

[illegible]

```

RESULT 5
US-08-665-716-2
; Sequence 2, Application US/08665716
; Patent No. 5789222
; GENERAL INFORMATION:
; APPLICANT: KELLY, ROSEMARIE
; APPLICANT: REGISTER, ELIZABETH A
; APPLICANT: MASUREKAR, PRAKASH S
; TITLE OF INVENTION: P5C REDUCTASE GENE FROM ZALERION
; TITLE OF INVENTION: ARBORICOLA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MERCK & CO., INC.
; STREET: 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: US
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,716
; FILING DATE: 23-JUN-1995

```

```

? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: KOESEN, ELLIOTT
? REGISTRATION NUMBER: 32,705
? REFERENCE/DOCKET NUMBER: 19453PVP
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 908-594-5493
? TELEFAX: 908-594-4720
? INFORMATION FOR SEQ. ID NO. 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 304 amino acids
? type: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
?
US-08-665-716-2

```

Query Match	18.88;	Score 294;	DB 1;	Length 304;
Best Local Similarity	29.98;	Pred. No. 6.2e-24;		
Matches	88;	Conservative	49;	Mismatches 119;
			Indels	38;
			Gaps	9

```

OY      1 MSVGFAGGOLAYNFYTAGILISAKRTIASSEPM-----LPT--VSAL-----42
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      11 LTLAVTGGGNGIALISGLTSLADEIHAPNSQSESTDETPSKLTKPIACRSKAGAKI 70
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      43 -----KGVNLTRSNKETV-KHSDVLEIAVKPHIIPILDETG-ADVOARHIWSCAA 93
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      71 KKALSPYTPKVIQTQSDNVATCREADVYLILCKEYPMAGILIEEGMDALGKLLISLA 13
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      94 GVTIISVYKKLMAQOPAP-----KYTRCMTNPVYVQEGATYATAGTHAL-VEDGQL 145
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      131 GVPABQITGYMIGKTPVNPPEKEGLCOVVRKAPNPAASGIREMTVIATISSPPLSTTSLSI 196
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      146 EQLMSSVGFCTEVEEDLIDAVTGLSGSPAYAFMALD--DGVKMGILPRRLATQLGAQA 203
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      191 TWIFRIGIDVQLPATMDASTALCGSSPAFFALILEAIDGAVAMGIPRAEAOAMAOT 256
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      204 LLGAARKMLLDBEQRHCOLKDWVCSPGAGTTHALHFLHESGSGRSLLINNAEASCI 257
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      251 MKGAGVLTSE-HPALLKDVTPPGGCTIGGLNVLEGGVGRGVARVAREATV 303
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 6  
US-08-605-150A-17  
Sequence 17 Application US/08605150A  
Patent No. 6103520  
GENERAL INFORMATION:  
APPLICANT: Topfer, Reinhard  
APPLICANT: Hausmann, Ludwig  
APPLICANT: Schell, Josef  
TITLE OF INVENTION: Glycerol-3-Phosphate Dehydrogenase  
TITLE OF INVENTION:  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klein & Szekeres  
STREET: 4199 Campus Drive, Suite 700  
CITY: Irvine  
STATE: CA  
COUNTRY: USA  
ZIP: 92715  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/605,150A  
FILING DATE: 01-MAR-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP94/02936  
FILING DATE: 02-SEP-1994  
PRIOR APPLICATION DATA:

```

1 APPLICATION NUMBER: DE P4329837.3
2 FILING DATE: 03-SEP-1993
3
4 ATTORNEY/AGENT INFORMATION:
5
6 NAME: Szekeres, Gabor B.
7 REGISTRATION NUMBER: 28,675
8 REFERENCE/DOCKET NUMBER: 542-04-PA
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: 714-854-5502
11 TELEFAX: 714-854-4897
12
13 INFORMATION FOR SEQ ID NO: 17:
14
15 SEQUENCE CHARACTERISTICS
16 LENGTH: 349 amino acids
17 TYPE: amino acid
18 TOPOLOGY: linear
19
20 US-08-605-150A-17

```

Query Match	6.5%;	Score 102;	DB 3;	Length 349;
Best Local Similarity	23.3%;	Pred. No. 0.0083;		
Matches	59;	Conservative	39;	Mismatches 95; Indels 60; Gaps 11

QY	59	SDVLFVAVPHIPIPTLDEICADVQARHIVSCAVTTSSEVKILMFOPAKVICM	118
Db	86	ADLIVEVPHQIGIKICQQLGHIKANTIGISILIKGVDEGPNGLKIS----	EVIGERL 1400
QY	119	MPVNVQESATVYANGTALVED-----	COLLEQILMSSVGF-CTVEE--- 1660
Db	141	GIPMSYILMANT---ASEVVAEKEFCETITGGCKDPAGOLILDKOTPNRITVQEDT	1366
QY	161	-DLIDVAVTSGSGPAVAFMMLADGCVKKGILPRILA-TQLGAOALLGAAMLLDSEOH	218
Db	197	VELICGLAKNIIVAVGAFGFC-----DELIGGDMDTKAAVIRLIGIMELIAFAKLF-----	2424
QY	219	QQLKDNVCSPGATTHALHTLESSEGFNSLLIMAVESCI--RREELMSADQKISPAAL	276
Db	243	-----CS-----GIWSSATFLFESCGVADLI-----TTCYGGRNRKVAEAFATGKTSIEOL	2877
QY	277	KKTLLDRVLTESP	289
Db	288	EKEMLNGOKLOGP	300

RESULT 7  
US-08-605-150A-6  
Sequence 6, Application US/08605150A  
Patent No. 6103520  
GENERAL INFORMATION:  
APPLICANT: Topfer, Reinhard  
APPLICANT: Hausmann, Ludwig  
APPLICANT: Scheell, Jozef  
TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROENASE  
TITLE OF INVENTION:  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klein & Szekeres  
STREET: 4199 Campus Drive, Suite 700  
CITY: Irvine  
STATE: CA  
COUNTRY: USA  
ZIP: 92715  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/605,150A  
FILING DATE: 01-MAR-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP94\02936  
FILING DATE: 02-SEP-1994  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P4329827.3  
 FILING DATE: 03-SEP-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Szekeres, Gabor L.  
 REGISTRATION NUMBER: 28,675  
 REFERENCE/DOCKET NUMBER: 542-04-PA  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 714-854-5502  
 TELEFAX: 714-854-4897  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 367 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-605-150A-6

Query Match 6.3%; Score 98; DB 3; Length 367;  
 Best Local Similarity 24.5%; Pred. No. 0.025;  
 Matches 70; Conservative 51; Mismatches 121; Indels 44; Gaps 15;

QY 43 KMGVNLTRSN--KETVHSDVLFPLAVKPH-----IIPFIIDEGADVQARHIV-----VSC 91  
 DB 83 KLGNVVAADPDLENVAKDANML-VFVTPHQFMEGICRKLVGKIOEGAQALSLIKMEYKM 141  
 QY 92 AAGVTISVEKKLMAFPAPKVRCTMTNPVVOEGAT-VYATGTHALVEDGOLLE---Q 147  
 DB 142 EGPCHMISSLISDLG-----INCCVLNGANIANEIAVEKFEATVFRENDIAEKWQ 195  
 QY 148 LMSVGECTEVEDLIDAVTGLSGSPAYAFMALDAD--GGVKKGLPRRLAI-QLGAQAL 204  
 DB 196 LFTSPYFVMSAVED---VEGVELCGTLKNIVAIAGFVDGLEMGNTKAAMRIGLREM 251  
 QY 205 LGAAKMLDSEQ-----HPCQLKD--NVCSPGATIHAIHFLESQGRSLINAVEASCI 257  
 DB 252 KAFKLLFPYSVKDTPFESCGVADLITTCIGGNRKVAEAFKNGNRS--FDLEAEML 309  
 QY 258 RTRELQSMADOEKISPAALKTKLLDRVKLE-SPYVSTLTTPSSPGKL 302  
 DB 310 RGOKLOGVSTAKEV-----YEVLRHKGWLELPFSTVHEISTGR 350

RESULT 8  
 US-08-605-150A-4  
 : Sequence 4, Application US/08605150A  
 : Patent No. 6103520  
 : GENERAL INFORMATION:  
 : APPLICANT: Topfer, Reinhard  
 : APPLICANT: Hausmann, Ludwig  
 : APPLICANT: Schell, Josef  
 : TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE  
 : NUMBER OF SEQUENCES: 17  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESS: Klein & Szekeres  
 : STREET: 4199 Campus Drive, Suite 700  
 : CITY: Irvine  
 : STATE: CA  
 : COUNTRY: USA  
 : ZIP: 92715  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/605,150A  
 : FILING DATE: 01-MAR-1996  
 : CLASSIFICATION: 800  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/EP94/02936  
 : FILING DATE: 02-SEP-1994

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE P4329827.3  
 FILING DATE: 03-SEP-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Szekeres, Gabor L.  
 REGISTRATION NUMBER: 28,675  
 REFERENCE/DOCKET NUMBER: 542-04-PA  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 714-854-5502  
 TELEFAX: 714-854-4897  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 381 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-605-150A-4

Query Match 6.3%; Score 98; DB 3; Length 381;  
 Best Local Similarity 24.5%; Pred. No. 0.026;  
 Matches 70; Conservative 51; Mismatches 121; Indels 44; Gaps 15;

QY 43 KMGVNLTRSN--KETVHSDVLFPLAVKPH-----IIPFIIDEGADVQARHIV-----VSC 91  
 DB 97 KLGNVVAADPDLENVAKDANML-VFVTPHQFMEGICRKLVGKIOEGAQALSLIKMEYKM 155  
 QY 92 AAGVTISVEKKLMAFPAPKVRCTMTNPVVOEGAT-VYATGTHALVEDGOLLE---Q 147  
 DB 156 EGPCHMISSLISDLG-----INCCVLNGANIANEIAVEKFEATVFRENDIAEKWQ 209  
 QY 148 LMSVGECTEVEDLIDAVTGLSGSPAYAFMALDAD--GGVKKGLPRRLAI-QLGAQAL 204  
 DB 210 LFTSPYFVMSAVED---VEGVELCGTLKNIVAIAGFVDGLEMGNTKAAMRIGLREM 265  
 QY 205 LGAAKMLDSEQ-----HPCQLKD--NVCSPGATIHAIHFLESQGRSLINAVEASCI 257  
 DB 266 KAFKLLFPYSVKDTPFESCGVADLITTCIGGNRKVAEAFKNGNRS--FDLEAEML 323  
 QY 258 RTRELQSMADOEKISPAALKTKLLDRVKLE-SPYVSTLTTPSSPGKL 302  
 DB 324 RGOKLOGVSTAKEV-----YEVLRHKGWLELPFSTVHEISTGR 364

RESULT 9  
 US-08-375-709-17  
 : Sequence 17, Application US/08375709  
 : Patent No. 5683898  
 : GENERAL INFORMATION:  
 : APPLICANT: YAZAWA, Kazunaga  
 : APPLICANT: YAMADA, Akiko  
 : APPLICANT: KATO, Seisho  
 : APPLICANT: KONDO, Kiyoaki  
 : TITLE OF INVENTION: Gene Coding For Elcosapentaenoic Acid  
 : TITLE OF INVENTION: Synthesizing Enzymes And Process for Production of  
 : NUMBER OF SEQUENCES: 19  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESS: Foley & Lardner  
 : STREET: 3000 K Street, N.W., Suite 500  
 : CITY: Washington  
 : STATE: D.C.  
 : COUNTRY: USA  
 : ZIP: 20007-5109  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/375,709  
 : FILING DATE: 20-JAN-1995  
 : CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/178,251  
FILING DATE: 14-MAY-1993  
PRIOR APPLICATION DATA: JP 4-147945  
FILING DATE: 15-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/150/AAOK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 543 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-375-709-17

Query Match 6.2%; Score 97.5; DB 1; Length 543;  
Best Local Similarity 22.5%; Pred. No. 0.053;  
Matches 71; Conservative 48; Mismatches 130; Indels 67; Gaps 13;

QY 8 AGOLAYRTAGIL-----SAHKTIASSPE-----MNLPTVSALRKMGVNLTRSNK- 53  
DB 119 AGILGSGFGAAGLIPSRKALINRIQALPNGPYMNLHSPSPALERGSVELFLKHV 178  
QY 54 ETVKHSDFLAVKPHIIFILDEIGADYQARHIVS-CAAGVTISSVEKKLMAFOPAPK 112  
DB 179 RTVEAS-AFLGLTPQIYVYRAGLSRDAGKVVGNKVIKAKSRTEVAEKEM--MPAP- 233  
QY 113 VIRGNTNPNVYVQEGATVATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGS 172  
DB 234 -----AKMLQKLVDDGSIYAEQEMELAOVLPMADDTTAEADSG--GH 272  
QY 173 GPAYAFMALDADGKVMGLPRRLAIQLGAQALIGAANKMLDSEQHPCOLKDNVCSFGAT 232  
DB 273 TDNRPLVTL-----LPTILALKEEIQ-----AKYQYDT---PIRVG---CGGSGVT 312  
QY 233 IHALHLESGGFRSLLINVAEASCI-----RTRELQSMADQEKISPAALKKTLIDRVKL 286  
DB 313 PDAALATFNMGAAYIVTGSINQACVEAGASDHTRKLLATTEMADVMAPAADMEKGVKL 372  
QY 287 ESPTVSTLTTPSSPGKL 302  
DB 373 QVVKRGTLFPMRANKL 388

RESULT 10  
US-08-752-929-17

Sequence 17, Application US/08752929  
Patent No. 5798259  
GENERAL INFORMATION:  
APPLICANT: YAZAWA, Kazunaga  
APPLICANT: YAMADA, Akiko  
APPLICANT: KATO, Seishi  
APPLICANT: KONDO, Kiyosi  
TITLE OF INVENTION: Gene Coding for Eicosapentaenoic Acid Synthesizing  
TITLE OF INVENTION: Enzymes and Process for Production of Eicosapentaenoic  
TITLE OF INVENTION: Acid  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/752,929  
FILING DATE: 20-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375,709  
FILING DATE: 20-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/178,251  
FILING DATE: 14-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-147945  
FILING DATE: 15-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/150/AAOK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 543 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-752-929-17

Query Match 6.2%; Score 97.5; DB 1; Length 543;  
Best Local Similarity 22.5%; Pred. No. 0.053;  
Matches 71; Conservative 48; Mismatches 130; Indels 67; Gaps 13;

QY 8 AGOLAYRTAGIL-----SAHKTIASSPE-----MNLPTVSALRKMGVNLTRSNK- 53  
DB 119 AGILGSGFGAAGLIPSRKALINRIQALPNGPYMNLHSPSPALERGSVELFLKHV 178  
QY 54 ETVKHSDFLAVKPHIIFILDEIGADYQARHIVS-CAAGVTISSVEKKLMAFOPAPK 112  
DB 179 RTVEAS-AFLGLTPQIYVYRAGLSRDAGKVVGNKVIKAKSRTEVAEKEM--MPAP- 233  
QY 113 VIRGNTNPNVYVQEGATVATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGS 172  
DB 234 -----AKMLQKLVDDGSIYAEQEMELAOVLPMADDTTAEADSG--GH 272  
QY 173 GPAYAFMALDADGKVMGLPRRLAIQLGAQALIGAANKMLDSEQHPCOLKDNVCSFGAT 232  
DB 273 TDNRPLVTL-----LPTILALKEEIQ-----AKYQYDT---PIRVG---CGGSGVT 312  
QY 233 IHALHLESGGFRSLLINVAEASCI-----RTRELQSMADQEKISPAALKKTLIDRVKL 286  
DB 313 PDAALATFNMGAAYIVTGSINQACVEAGASDHTRKLLATTEMADVMAPAADMEKGVKL 372  
QY 287 ESPTVSTLTTPSSPGKL 302  
DB 373 QVVKRGTLFPMRANKL 388

RESULT 11

US-09-090-793-10  
Sequence 10, Application US/09090793  
Patent No. 6140486  
GENERAL INFORMATION:  
APPLICANT: Calgene, LLC  
TITLE OF INVENTION: Production of polyunsaturated fatty acids by expression  
TITLE OF INVENTION: of polyketide-like synthesis genes in plants  
FILE REFERENCE: CGNE.131.0105  
CURRENT APPLICATION NUMBER: US/09/090,793

CURRENT FILING DATE: 1998-06-04  
 EARLIER APPLICATION NUMBER: 60/048, 650  
 EARLIER FILING DATE: 1997-06-04  
 NUMBER OF SEQ ID NOS: 66  
 SOFTWARE: Patent In Ver. 2.0  
 SEQ ID NO 10  
 LENGTH: 543  
 TYPE: PRT  
 ORGANISM: Shewanella putrefaciens  
 US-09-090-793-10

Query Match 6.2%; Score 97.5; DB 4; Length 543;  
 Best Local Similarity 22.5%; Pred. No. 0.053;  
 Matches 71; Conservative 48; Mismatches 130; Indels 67; Gaps 13;

QY 8 ACOLAVRTAACIL-----SAHKTIASSPE-----MNLPTYSALKKMGVNLTRSKK- 53  
 DB 119 AGILGSGFAGLIPSRVEALINRIQALPNGPYMNLHSPSEPALRGSEVLEFLKHKV 178  
 QY 54 ETVKHSDFLAVKPHIIPILDEIGADVQARHIVS--CAAGVTISSEVKLMAFOPAPK 112  
 DB 179 RTVEAS--AFGLTPOIYVYRAAGSRDAQGVVGNKVIAKVSNTVEAEKFM--MPAP- 233  
 QY 113 VIRCMTPPVVVEGATVATGTHALVEDGQLLEQLMSSVGEVEEDLIDAVTGLSGS 172  
 DB 234 -----AKMLQKLVDDGSTTAEOMELAQVPMADITAEADSG--GH 272  
 QY 173 GAVYAFMALDAGGVKMLPRRLATQLGAQALLGAKKMLDSEHQPOLKONVCSPGAT 232  
 DB 273 TDNRPLVTL-----LPTILAKEEIQ-----AKYQYDT--PIVAG--CGGAGVT 312  
 QY 233 IHALHLESGGFRSLINAVEASCI-----RTRELQSMADOEKISPAALKKTLIDRVKL 286  
 DB 313 PRAALATFMGAIYVITGSGINACVACVAGASDHTKRLATTEMADVTMAPADMEFMGVL 372  
 QY 287 ESPVSTLTSPSSPKL 302  
 DB 373 QVKKRGTLEPMRANKL 388

RESULT 12  
 US-08-726-012B-2  
 Sequence 2, Application US/08726012B  
 Patent No. 5952190  
 GENERAL INFORMATION:  
 APPLICANT: Hans Joenje, et al.  
 TITLE OF INVENTION: CDNA FOR FANCONI ANEMIA COMPLEMENTATION GROUP A  
 NUMBER OF SEQUENCES: 18  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Klarquist Sparkman Campbell Leigh & Whinston, LLP  
 STREET: One World Trade Center, Suite 1600, 121 S.W. Salmon Street  
 CITY: Portland  
 STATE: OR  
 COUNTRY: USA  
 ZIP: 97204-2968  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Disk, 3.5-inch  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: MS DOS  
 SOFTWARE: WordPerfect 5.1+, ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/726,012B  
 FILING DATE: 10/04/96  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Richard J. Polley  
 REGISTRATION NUMBER: 28,107  
 REFERENCE/DOCKET NUMBER: 3812-45520/RJP/DJE  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (503) 226-7391  
 TELEFAX: (503) 228-9446  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1455  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-08-726-012B-2

Query Match 6.1%; Score 95.5; DB 2; Length 1455;  
 Best Local Similarity 22.3%; Pred. No. 0.43;  
 Matches 79; Conservative 42; Mismatches 130; Indels 103; Gaps 14;

QY 19 GILSAHKTIASSPEMNLPTYSALKKMGVNLTRSKKTVKHSKY----- 61  
 DB 189 GIVSLQELSHEDMAVSGMELRNL--CCICEQMEASCOIADVARMLSDVQMFVLRGP 247  
 QY 62 -----FLAVKPHIIP-----FLDEIGADVQ--ARHIVSCAGVTISSEV 101  
 DB 248 OKNSDLRRVEPEKMPQVTVDLQRLIFALDALLAGVDESSTHKIVRCMFVFGHTL 307  
 QY 102 KKLMAEQAPARK-----VIRCMTPPVV-----VQEGATVATGTHALVED- 141  
 DB 308 GSVISTDPLKRFPSHTLTQILTHSPVLKASDAVQOMQREMSFARTHTPLTSLYRRLVMS 367  
 QY 142 -----GOLLQOL-----MSSVGFQ--TEVEEDLIDAVTGLSGSPYAFMALD 182  
 DB 368 ABEVLGHQLEVLTEQVHMQRVLSFVSALVCFPEAQQLLEDVVARLM-----AQAFESQ 423  
 QY 183 ADGVKMGRLPRRLATQLGAQALLGAKKMLDS-----EQHPQOLKONV- 226  
 DB 424 LDSVTAFLVLRQALBGPAPFLSYADWFKASFGSTRGVHGCCKKMLVFLTFLSELVPE 483  
 QY 227 -SPGATIIHALH-FLBSGFRSLINAVEASCIPTREL-----OSMADOEKISPA 274  
 DB 484 ESPRYQVHILHPLPVSRYSLTLTYISLAKTRLADLAKVSTENMGLEYEDLSA 537

RESULT 13  
 US-08-605-150A-8  
 Sequence 8, Application US/08605150A  
 Patent No. 6103520  
 GENERAL INFORMATION:  
 APPLICANT: Topfer, Reinhard  
 APPLICANT: Hausmann, Ludwig  
 APPLICANT: Schell, Jozeff  
 TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE  
 NUMBER OF SEQUENCES: 17  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Klein & Szekeres  
 STREET: 4199 Campus Drive, Suite 700  
 CITY: Irvine  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 92715  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/605,150A  
 FILING DATE: 01-MAR-1996  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/EP94/02936  
 FILING DATE: 02-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE P4329827.3  
 FILING DATE: 03-SEP-1993

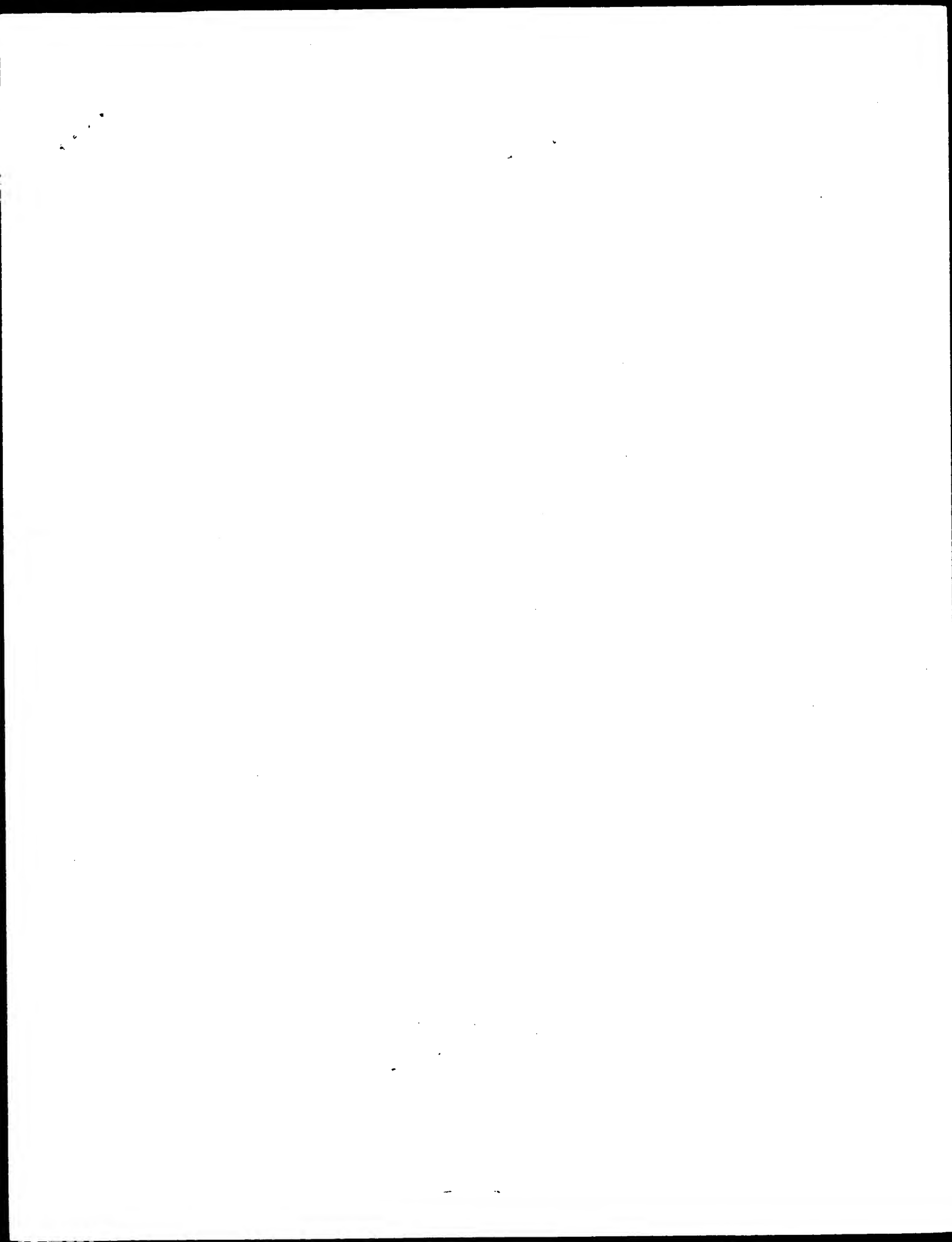
; PRIOR APPLICATION DATA:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:54:15 ; Search time 12.98 Seconds

(without alignments)  
886.961 Million cell updates/sec

Title: US-09-912-717-1

Perfect score: 1567  
Sequence: 1 MSVFGTGAQLAYRTAGI.....TPSSPGKLLTRSLALGKKMD 314

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1306	83.3	319	1	PROG_HUMAN
2	512.5	32.7	276	1	PROG_ARATH
3	501.5	32.0	274	1	PROG_SOYBN
4	479	30.6	273	1	PROG_PEA
5	476	30.4	278	1	PROG_ACTYH
6	421	26.9	267	1	PROG_SYNY3
7	402.5	25.7	269	1	PROG_ECOLI
8	377	24.1	272	1	PROG_PSEAE
9	370	23.6	270	1	PROG_CORGL
10	345	22.0	270	1	PROG_BACSU
11	326.5	20.8	294	1	PROG_MYCLE
12	324.5	20.7	278	1	PROG_BACSU
13	323.5	20.6	261	1	PROG_THERM
14	306.5	19.6	295	1	PROG_MYCTU
15	301.5	19.2	311	1	PROG_NEUCR
16	297	19.0	265	1	PROG_AOUAE
17	296	18.9	320	1	PROG_ZALAR
18	281.5	18.0	286	1	PROG_YEAST
19	279	17.8	293	1	PROG_TREPA
20	266	17.0	299	1	PROG_CAEBL
21	264	16.8	271	1	PROG_HAETN
22	252.5	16.1	278	1	PROG_VIBAL
23	193.5	12.3	257	1	PROG_HELPY
24	179.5	11.5	257	1	PROG_HELPY
25	163.5	10.4	273	1	CME4_BACSU
26	109	7.0	599	1	CENB_MOUSE
27	106	6.8	348	1	GPDA_RABIT
28	104	6.6	272	1	YKFA_BACSU
29	104	6.6	451	1	YEVJ_METTF
30	102	6.5	251	1	PROG_METSM
31	102	6.5	348	1	GPDA_MOUSE
32	101	6.4	599	1	CENB_HUMAN
33	100	6.4	348	1	GPDA_RAT

34	97	6.2	272	1	THIM_BACSU
35	97	6.2	348	1	GPDA_HUMAN
36	97	6.2	351	1	GPDA_FUGRU
37	95.5	6.1	1455	1	FACA_HUMAN
38	91.5	5.8	346	1	ARCC_BACSU
39	91.5	5.8	347	1	ID12_ERKHE
40	91	5.8	790	1	SYFB_CHIMU
41	90	5.7	397	1	TRB1_AOUAE
42	89.5	5.7	371	1	MAIK_ECOLI
43	89.5	5.7	569	1	AMP2_LYCES
44	89.5	5.7	2541	1	TALI_HUMAN
45	89	5.7	906	1	KBP2_CHICK

## ALIGNMENTS

RESULT	ID	PROG_HUMAN	STANDARD	PRT	319 AA.
AC	P33322				
DT	01-OCR-1993	(Rel. 27, Created)			
DT	01-OCR-1993	(Rel. 27, Last sequence update)			
DT	20-AUG-2001	(Rel. 40, Last annotation update)			
DE	PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).				
GN	PYCRL				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92112821; PubMed=1730675;				
RA	Dougherty K.M., Brandtiss M.C., Valle D.;				
RT	*Cloning human pyrrolidine-5-carboxylate reductase cDNA by				
RT	complementation in Saccharomyces cerevisiae.*				
RL	J. Biol. Chem. 267:871-875(1992).				
CC	-1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = L-PYRROLINE-5-CARBOXYLATE + NAD(P)H.				
CC	-1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.				
CC	-1- SUBUNIT: HOMODECAMER OR HOMODODECAMER.				
CC	-1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
DR	EMBL: M77836; AAA36407.1; -.				
DR	PIR: A41770; A41770.				
DR	MIM: 179035; -.				
DR	InterPro: IPR000304; P5CR.				
DR	Pfam: PF01089; P5CR.1.				
DR	PROSITE: PS00521; P5CR.1.				
KW	Oxidoreductase; proline biosynthesis; NADP.				
SQ	SEQUENCE 319 AA; 33374 MW; F5E74B5BDF475EF CRC64;				

Query Match 83.3%; Score 1306; DB 1; Length 319;  
Best Local Similarity 82.7%; Pred. No. 1.5e-89;  
Matches 263; Conservative 27; Mismatches 22; Indels 6; Gaps 2;

QY	1	MSVFGTGAQLAYR---FTAGILSAHKITASSPENMLPTVSALRKMGVLTIRSNKETV 56
DB	1	MSVFGTGAQLAFLALGKFTAGVLAHAKIMASSPMDLTVSALRKMGVLTIRPNKETV 60
QY	57	KHSVFLFLAVKPHIPIPIIDIGADVAGARHIVVSCAGVYTISSVEKKLMAFORAPVIRC 116
DB	61	QHSDFLFLAVKPHIPIPIIDIGADIEDRIHVSCAGVYTISSIEKKLSAFRAPVIRC 120

```

QY 117 MNTPEVVOEGANVATGTHALVEDEGOLLEOLMSSVGFCTEVEEDLIDAVTGLSGSPAY 176
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 MNTPEVVOEGANVATGTHALVEDEGOLLEOLMSSVGFCTEVEEDLIDAVTGLSGSPAY 180
QY 177 AFMALD--ADGGVKGMLPRRLAIGLGAQALLGAARKMLDSEOHPCOLKDNVCSPGATIH 234
    || ||| |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 AFDALDALADGGVKGMLPRRLAIGLGAQALLGAARKMLDSEOHPCOLKDNVCSPGATIH 240
QY 235 ALHFLDSGFRSLINAVASCIPTRELOSMADEKISPAALKTLLDRKLESFTVSTL 294
    || ||| |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 ALHFLDSGFRSLINAVASCIPTRELOSMADEKISPAALKTLLDRKLESFTVSTL 300
QY 295 TPSSPGKLTLSLALGK 312
    || ||| ||| ||| ||
Db 301 SPGGHKLPRSLAPAGK 318

RESULT 2
PROC_ARATH STANDARD; PRT: 276 AA.
ID PROC_ARATH STANDARD; PRT: 276 AA.
AC P54904:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
GN PROCI OR AT5G14800 OR T9L3.100.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94294559; PubMed=8022935;
RA Verrijgen N., Villarroel R., van Montagu M.;
RT "Osmoregulation of a pyrroline-5-carboxylate reductase gene in
RT Arabidopsis thaliana."
RL Plant Physiol. 103:771-781(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG ERECTA;
RA Verrijgen N., Villarroel R., Hua X., van Montagu M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=21016721; PubMed=11130714;
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
RA Nakazaki N., Naro K., Okumura S., Shino S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA Leonard S., Meyer R., Milvanev E., Ozerky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Wolam A., Yokum M., Bell M., Dedina N.,
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vill D., Baker J.,
RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
RA Marijnsen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Entlan K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McCullagh B., Robben J., Gijmenez B., Zimmermann W.,
RA Ransperger U., Wedler H., Balke K., Wedler E., Peters S.,
RA van Staveren M., Dirke W., Mooljman P., Klein Lankhorst R.,
RA Weitzenecker T., Bothe G., Rose M., Hauf J., Bernies S., Hempel S.,
RA Feilspausch M., Lamberth S., Villarroel R., Gietlen J., Arlies W.,
RA Behts O., Lemke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
RA Scheller C., Zaccaria P., Mewes H.-W., Beyer M., Franz P.F.;
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis

```

```

RT thaliana."
RL Nature 408:823-826(2000).
CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = L-PYRROLINE-5-
CC CARBOXYLATE + NAD(P)H.
CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M76538; AAA61346.1; -.
DR EMBL: Y08951; CAAT0148.1; -.
DR EMBL: AL391149; CAC01879.1; -.
DR HSP: P80028; 1TOF.
DR InterPro: IPR000304; P5CR.
DR Pfam: PF01089; P5CR.1.
DR PROSITE: PS00521; P5CR.1.
KW Oxidoreductase; Proline biosynthesis; NADP.
SQ SEQUENCE 276 AA; 28624 MW; B577A01C92A3A28B CRC64;

Query Match 32.7%; Score 512.5; DB 1; Length 276;
Best Local Similarity 42.5%; Pred. No. 4,8e-31;
Matches 113; Conservative 52; Mismatches 90; Indels 11; Gaps 4;

QY 3 VGFILGAGOLAYRF---TAAGILSAHKITLASSPEMNLPFTSALRKMGVNTLRNKKETVKH 58
    |||||||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13 VGFILGAGKMAESTIRGVYASVLPNR-ICTAHSNLRNDVEESGVNFTSSEVYKE 71
QY 59 SDVILFVAKPHIIPILDEIGADVQARHIVSCAGVYISVKKLMAFQAPADRVRCMT 118
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 72 SDVIFSVKPVYKAVKATELKSRLKNIIVSAAGIKLNDIE---WSGDQRFIVMP 127
QY 119 MNTPEVVOEGANVATGTHALVEDEGOLLEOLMSSVGFCTEVEEDLIDAVTGLSGSPAY 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 MTPAAGEAASVMSLGGATEEDGATVAMLPFGAGKTLKADKMFDPVAVTGLSGSPAY 187
QY 179 MALD--ADGGVKGMLPRRLAIGLGAQALLGAARKMLDSEOHPCOLKDNVCSPGATIH 236
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 188 LAIALADGGVAGNAPRLRALSLASQTVLGAATVNSKTKHPVLKDDVTPSGTITAGV 247
QY 237 HFLDSGFRSLINAVASCIPTREL 262
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 HELESGFRSLINAVASCIPTREL 273

RESULT 3
PROC_SOYBN STANDARD; PRT: 274 AA.
ID PROC_SOYBN STANDARD; PRT: 274 AA.
AC P17817:
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90340278; PubMed=2199815;
RA Delauney A.J., Verma D.P.S.;
RT "A soybean gene encoding delta 1-pyrroline-5-carboxylate reductase
RT was isolated by functional complementation in Escherichia coli and is
RT found to be osmoregulated."

```

RL Mol. Genet. 221:299-305(1990).  
CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = 1-PYRROLINE-5-CARBOXYLATE + NAD(P)H.  
CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES, BUT MOSTLY IN NODULES.  
CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X16352; CAA34401.1; -  
CC PIR: S10186; S10186.  
CC InterPro: IPR000304; P5CR.  
CC Pfam: PF01089; P5CR; 1.  
CC PROSITE: PS00521; P5CR; 1.  
CC Oxidoreductase; Proline biosynthesis; NADP.  
CC SEQUENCE 274 AA: 28586 MW: 933CFCFD07598B63 CRC64;  
SQ

Query Match 32.0%; Score 501.5; DB 1; Length 274;  
Best Local Similarity 41.6%; Pred. No. 3.1e-30;  
Matches 111: Conservative 51; Mismatches 94; Indels 11; Gaps 4;

QY 2 SVGTGAGQLAYR---TAAGILSAHKIIASSPEMLPTVSALKMGVNTLRNKETV 57  
DB 12 TLGTGAGKMAESIRKAGVRSGLVPPSR-IRTAHVHNLARKGAFESFVTVLPSNDVVR 70  
DB 58 HSDVFLAVKPHIIFILDEIGADVQARHIYVSCAGYTISSVEKKLMAFPAPAKVIR 117  
DB 71 ESDVYVLYKPKDLYVDVYKLPPLTKKLLVSAVAGTKL---KDQEMGHERFTR 126  
QY 118 TPTPVVOGCAVYVYVGTALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSPA 177  
DB 127 PTPPAVGAQASVMSLGSATBEDGNTIAQLFGSIGKIMKAEKFDIAITGLSSGPA 186  
QY 178 FMALD-ADGVKMGILPRLAIQLGAQALLGAKMLDSEQHPCQLKDNVCSPGATT 235  
DB 187 YLAIFALADGVAAGLPRLSLASQVTLGAASVWSOTGKHPCQLKDDVISPCTITG 246  
QY 236 LHFLESGGRSLILINAVASCIPTREL 262  
DB 247 IHELENGGFRGTLMAVVAARSRREL 273  
RESULT 4  
PROC\_PEA STANDARD: PRT: 273 AA.  
ID 004708;  
AC 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).  
GN PROC.  
OS Pisum sativum (Garden pea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;  
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
OX NCBI\_TaxID=3888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Williamson C.L., Slocum R.D.;  
RT Molecular cloning and evidence for osmoregulation of the F1-  
RT pyrroline-5-carboxylate reductase (proc) gene in pea (Pisum sativum  
RT L.)";  
RL Plant Physiol. 100:1464-1470(1992).

CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = 1-PYRROLINE-5-CARBOXYLATE + NAD(P)H.  
CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X62842; CAA44646.1; -  
CC InterPro: IPR000304; P5CR.  
CC Pfam: PF01089; P5CR; 1.  
CC PROSITE: PS00521; P5CR; 1.  
CC Oxidoreductase; Proline biosynthesis; NADP.  
CC SEQUENCE 273 AA: 28237 MW: D719F52935966588 CRC64;  
SQ

Query Match 30.6%; Score 479; DB 1; Length 273;  
Best Local Similarity 40.1%; Pred. No. 1.4e-28;  
Matches 108: Conservative 55; Mismatches 90; Indels 16; Gaps 5;

QY 2 SVGTGAGQLAYR---TAAGILSAHKIIA--SPEMLPTVSALKMGVNTLRNKET 55  
DB 12 TLGTGAGKMAESIRKAGVRSGLVPPSRIRTAHVHNLARKGAFESFVTVLPSNDV 67  
QY 56 VKHSDVFLAVKPHIIFILDEIGADVQARHIYVSCAGYTISSVEKKLMAFPAPAKVIR 115  
DB 68 VRASNVYVFSKPKDLYVDVYKLPPLTKKLLVSAVAGTKL---KDQEMGHERFTR 123  
QY 116 TPTPVVOGCAVYVYVGTALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSPA 175  
DB 124 VMPPTPAVGAQASVMSLGSATBEDGNTIAQLFGSIGKIMKAEKFDIAITGLSSGPA 183  
QY 176 FMALD-ADGVKMGILPRLAIQLGAQALLGAKMLDSEQHPCQLKDNVCSPGATT 233  
DB 184 YLAIFALADGVAAGLPRLSLASQVTLGAASVWSOTGKHPCQLKDDVISPCTITG 243  
QY 234 LHFLESGGRSLILINAVASCIPTREL 262  
DB 244 AGVHELENGGFRGTLMAVVAARSRREL 272

RESULT 5  
PROC\_ACTCH STANDARD: PRT: 278 AA.  
ID 004016;  
AC 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).  
GN PROC.  
OS Actinidia chinensis (kiwifruit) (Yangtze).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;  
OC Asceridae; Ericales; Actinidiaceae; Actinidia.  
OX NCBI\_TaxID=3625;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA TISSUD-AXillary bud.  
RC Walton E.F., Podivinsky E., Wu R.M., Reynolds P.H.S., Young L.W.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = 1-PYRROLINE-5-CARBOXYLATE + NAD(P)H.  
CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

RP SEQUENCE FROM N.A.  
BA Duncan M Allen F Araujo B Amaral A M Chung F Davis K



Db 237 EAAKSFQANGFEALVEQALNNAORSALAEQLO 272

# RESULT 9

PROC\_CONGL STANDARD; PRT; 270 AA.

ID PROC\_CONGL

AC P46540;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE PYRROLIN-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).

OS Corynebacterium glutamicum (Brevibacterium flavum).

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;

OC Corynebacterium.

OX NCBI\_TaxID=1718;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 17965 / Melassecola;

RA MEDLINE=96345604; PubMed=8755867;

RT "Mutations in the Corynebacterium glutamicum proline biosynthetic

pathway: a natural bypass of the pro step.";

RT J. Bacteriol. 178:4412-4419(1996).

CC -1 CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = L-PYRROLIN-5-

CARBOXYLATE + NAD(P)H.

CC -1 PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.

CC -1 SIMILARITY: BELONGS TO THE PYRROLIN-5-CARBOXYLATE REDUCTASE

FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: U31225; AAC4172.1; -

DR InterPro: IPR000304; P5CR.

DR Pfam: PF01089; P5CR; 1.

DR PROSITE: PS00521; P5CR; 1.

KW Oxidoreductase; Proline biosynthesis; NADP.

SEQUENCE 270 AA: 28223 MW; 90150E233D94158F CRC64;

Query Match 23.6%; Score 370; DB 1; Length 270;

Best Local Similarity 33.8%; Pred. No. 1.6e-20;

Matches 93; Conservative 56; Mismatches 102; Indels 24; Gaps 8;

2 SVCFIAGQLAIRFTAGLSAHTKIASPEKMLPVSLAKRMG-----VULTRSN 52

Db 3 TAVIGGGIG- EALVSGILIAN---NMPO-NIRVTNSEERGGELRDYGIIMT-DN 55

53 KEVKGSDVLAFAKPHIIPFIDELIGADV---QARHIVSCAGVTSVEKTIAROP 109

Db 56 SQADADADVFLCYKKEFLVEISETGLDNNNSQSVVSAAGSTIAMESASAGLP 115

110 AKVIRICMTNTPVVQEGATVATGTHALVEDQLEIQLMSSVGFCTEVEEDLDAVTGL 169

Db 116 ---VVHVMNTPTLVGKGSITWKGRYVDAEQLGVKDLSTGVDLVEAESDIDAVTAM 172

170 SSGSPAYVAMALDA--DGVKMGKLPRLAIOGAQALLGAAMLLDSEHQPOLDNVCS 227

Db 173 SSGSPAYVLFVLEALIEAVNLGLPRATAKKLAVASFECAATMKTGKPESELRAVSS 232

228 PGGATTHALFLESFGFRSLLINAVASCITRREL 262

Db 243 PACTVTAIRLEESGIRGAFYRAQACADRSSEL 267

RESULT 10

PROH\_BACSU STANDARD; PRT; 270 AA.

ID PROH\_BACSU

AC P14363; 031828;

DT 01-JAN-1990 (Rel. 13, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE PYRROLIN-5-CARBOXYLATE REDUCTASE HOMOLOG 1.

GN YOKE OR PROC.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI\_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;

RT Submitted (NOV-1997) to the EMBL/GenBank/DBD databases.

RN [2]

RP SEQUENCE OF 1-256 FROM N.A.

RC STRAIN=W23; AND 168;

RA MEDLINE=91192601; PubMed=1849493;

RT "Variations and coding features of the sequence spanning the

replication terminus of Bacillus subtilis 168 and W23 chromosomes.";

RT Gene 98:107-112(1991).

RN [3]

RP SEQUENCE OF 1-200 FROM N.A.

RC STRAIN=W23;

RA MEDLINE=89155440; PubMed=2493444;

RT "DNA and protein sequence conservation at the replication terminus in

Bacillus subtilis 168 and W23.";

RT J. Bacteriol. 171:1402-1408(1989).

CC -1 FUNCTION: NOT KNOWN, BUT COULD BE INVOLVED IN GLUTAMINE

BIOSYNTHESIS.

CC -1 SIMILARITY: BELONGS TO THE PYRROLIN-5-CARBOXYLATE REDUCTASE

FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: Z99114; CAB13741.1; -

DR EMBL: M24523; AAA22722.1; -

DR PIR: B32807; B32807.

DR Subtilist; Bg11049; YOKE.

DR InterPro: IPR000304; P5CR.

DR Pfam: PF01089; P5CR; 1.

KW PROSITE: PS00521; P5CR; 1.

KW Hypothetical protein; Oxidoreductase; Complete proteome.

FT VARIANT 58 58 T -> A (IN STRAIN W23).

FT VARIANT 70 70 L -> S (IN STRAIN W23).

FT VARIANT 74 74 N -> I (IN STRAIN W23).

FT VARIANT 93 93 L -> S (IN STRAIN W23).

FT VARIANT 100 100 S -> T (IN STRAIN W23).

FT VARIANT 127 127 E -> Q (IN STRAIN W23).

FT VARIANT 159 159 K -> Q (IN STRAIN W23).

FT VARIANT 236 236 H -> Q (IN STRAIN W23).

FT VARIANT 239 239 I -> V (IN STRAIN W23).

FT CONFLICT 256 256 Q -> E (IN REF. 2).

SEQUENCE 270 AA: 29402 MW; C7B9E1680B7F9A7 CRC64;

Query Match 22.0%; Score 345; DB 1; Length 270;

Best Local Similarity 34.6%; Pred. No. 1.1e-18;

Matches 88; Conservative 53; Mismatches 101; Indels 12; Gaps 6;

3 VCFIAGQLAIRFTAGLSAHTKIASPEKMLPVSLAKRMG-----VULTRSN 57



```

Db      19 VAFIGGSGMA -EEMISGIYVANKIPKONICVTRNSMTERLETELQYGIKALPNQCLIE 77
Oy      58 HSGVLEAVKPHIIPRIIDELICADVDQARHIVVSCAGVITSSVEKLMAPQAPKVIKRC 117
Db      78 DMDVLTILAKPKDAENALSLKSRIOPHOLITSLVLAGITTSFEOSLNEQP---VVRVM 134
Oy      118 TMTPVVQECATVYANGTATLVEDGQLEEQLMSSVFCFTEVEEDLIDAVYGLSGSPAVA 177
Db      135 PNTSSMIGASATVAINLKGKVSBDLKLALBLLGCMEEYVITQNMQNDIFLIGSGSPATV 194
Oy      178 FNAUD--ADGCVKMGLPRLAIOGLAQVALLGAACKMLLDSEQHPCQKIDNVCSFGATTHA 235
Db      195 YLMEPIENTGEEAGIDKOLSRSIGAQTLLGAAKMLETGEHPEILRDNTSPNGTTAAG 254
Oy      236 LHFV-ESGGRSL 248
Db      255 LQALKKSGGEKRF 268

RESULT 11
PROC_MYCLE
ID      PROC_MYCLE      STANDARD:      PRT:      294 AA.
AC      P46725:
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      20-AUG-2001 (Rel. 40, Last annotation update)
DE      PYRROLATE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
GN      PROC OR M12430 OR B2168_C2_211.
OS      Mycobacterium lepreae.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1769;
      (1)
RA      SEQUENCE FROM N.A.
RA      Smith D.R., Robison K.;
RL      Submitted (MAR-1994) to the EMBL/Genbank/DBSJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=TN;
RX      MEDLINE=21128732; PubMed=11234002;
RA      Cole S.T., Elgmier K., Parkhill J., James K.D., Thomson N.R.,
RA      Wheeler P.R., Honore N.K., Garnier T., Churcher C., Harris D.,
RA      Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R.M., Devlin K., Duthoy S., Feltingworth T., Fraser A., Hamlin N.,
RA      Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moulé S.,
RA      Murphy L., Oliver K., Quail M.A., Rajandream M.-A., Rutherford K.M.,
RA      Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA      Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA      Barrall B.G.;
RT      "Massive gene decay in the leprosy bacillus.";
RL      Nature 409:1007-1011(2001).
CC      -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = L-PYRROLINE-5-
CC      CARBOXYLATE + NAD(P)H.
CC      -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
CC      -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U000018; AAA17233.1; -.
DR      EMBL; AL583925; CAC31947.1; -.
DR      Lepronia; ML2430; -.
DR      InterPro; IPR000304; P5CR.
DR      Pfam; PF01089; P5CR.1.
DR      PROSITE; PS00521; P5CR.1.
KW      Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
KW      SEQUENCE 294 AA; 30237 MW; EA8606C9CB8B69DB CQC64;

```

Query Match	20.8%	Score 326.5	DB 1	Length 294
Best Local Similarity	36.2%	Pred. No. 2,8e-17		
Matches	88	Conservative	45	Mismatches 77
				Indels 33
				Gaps 7

RESULT	12	
PROJ_BACSU	STANDARD:	PRT; 278 AA.
AC	PSA552;	
DT	01-OCT-1996 (Rel. 34, Created)	
DT	01-OCT-1996 (Rel. 34, Last sequence update)	
DT	20-AUG-2001 (Rel. 40, Last annotation update)	
DE	PYRROLINE-5-CARBOXYLATE REDUCTASE HOMOLOG 2.	
GN	YOJO.	
OS	Bacillus subtilis.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;	
CC	Bacillus/Staphylococcus group; Bacillus.	
OX	NCBI_TaxID=1423;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=168 / JH642.	
RA	Kobayashi Y., Mizuno M., Masuda S., Takamaru K., Hosono S.,	
RL	Sato T., Takeuchi M.;	
CC	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.	
CC	-1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	-----	
DR	EMBL; D84432; BAA12621.1; -;	
DR	EMBL; Z59116; CAB14312.1; -;	
DR	Subtilist; Bg11744; YJ2JO.	
DR	InterPro; IPR000304; P5CR.	
DR	Pfam; PF01089; P5CR.1.	
DR	PROSITE; PS00521; P5CR.1.	
DR	Hypothetical protein; oxidoreductase; Complete proteome.	
DR	SEQUENCE 278 AA; 30396 MW; 48695B52E55345E CRC64;	

```
Query Match      20.7%   Score 324.5; DB 1, Length 278;
Best Local Similarity 29.9%   Pred. NO. 3,7e-17;
Matches 88; Conservative 54; Mismatches 113; Indels 39; Gaps 7;

OY    3 VGRIGAGQLAIR-----FTAAGLISLAHKKIASSPEKNILPTYSALKRKGKVNLTIRSKKETVHK 58
      :|::|||:-|          |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```

Query Match 20.6%; Score 323.5; DB 1; Length 261;
Best Local Similarity 33.9%; Pred. No. 4,1e-17;
Matches 93; Conservative 45; Mismatches 107; Indels 29; Gaps 11;

QY 1 MSVGIGAGOLAYRPTAA---GILSAHK-IIASPEENIPTVSALRRMGVN-ITRSNK 53
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 41 MLRAFGGKMGORSILKGLERGFRLPEEYGLGCRPEKRS---ETAEFGVRLPTRAD- 56
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 54 EIVKHSDFVFLAKVKKHIIIFILDEIGADYQAHIV--YSCAGVITSSYEKTLMFQRPAP 111
   : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```

Db      57 -LGMAEVVLAVQPRDPRLAPEL-----AHNHLGIVISIMAGISTSVLARRL-----DNR   105
Oy      112 RVRCMTMPVPVVGEGATVYATGTTHAL-VEDGQLLEOLMSSVGCETEVEEDLDIAYTGLS   170
          :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      106 RVRMRPMPLAIVIGSGESTALTALKAREAREADAFARALLFAFVGVTEPEHLPFLFTGMS    165
Oy      171 GGGPAYAAFMALD--ADGGVKMGDLPRRLATOAGAOLLGAAMKMLDSRQHPCOLDKNYCSP     228
          ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      166 ASAPPYLVAVALDALDACGYKMGRPALRALRIADALAATGELL--KGHRPQGVDEVASP     223
Oy      229 GGATIHAIHFLESGGFSRLINAVEASCIRTEL    262
          ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      224 GSTTIHGHALERAVRAAFYEAVEEAATRQHEL    257

RESULT   14
PROC_MYCTU        STANDARD;              PRT;           295 AA.
ID            AC         Q11141;
DT            DT          01-OCT-1996 (Rel. 34, Created)
DT            DT          01-OCT-1996 (Rel. 34, Last sequence update)
DT            DT          20-AUG-2001 (Rel. 40, Last annotation update)
DE            DE          PYROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (5SCR) (5SC REDUCTASE).
GN            GN          PROC OR RV0500 OR MT0520 OR MTCY20G9.26.
OS            OS          Mycobacterium tuberculosis.
OC            OC          Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OX            OX          Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium,
RN            RN          NCBI_TaxId=1773;
RP            RP          [1]
RM            RM          SEQUENCE FROM N.A.
RX            RX          MEDLINE=96295987; PubMed=9634230;
RA            RA          Cole S.T., Brosch R., Parkhill J., Garner T., Churcher C., Harris D.,
RA            RA          Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA            RA          Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA            RA          Davies R., Dellin K., Fellwell T., Gentile S., Hamlin N., Holroyd S.,
RA            RA          Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA            RA          Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA            RA          Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA            RA          Sutton J.E., Taylor K., Whitehead S., Barrett B.G.;
RT            RT          "Deciphering the biology of Mycobacterium tuberculosis from the
RT            RT          complete genome sequence.";
RL            RL          Nature 393:537-544(1998).
RN            RN          [2]
RP            RP          SEQUENCE FROM N.A.
RC            RC          STRAIN=CDC 1551 / Oshkosh;
RA            RA          Fleetschman R.D., Allard D., Eisen J.A., Carpenter L., White O.,
RA            RA          Peterson S.J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA            RA          Kolonay J.F., Nelson W.C., Umayang I.A., Ermolaeva M.D., Salzberg S.L.,
RA            RA          Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RT            RT          Bishop W.;
RL            RL          "Whole genome comparison of Mycobacterium tuberculosis clinical and
          laboratory strains."
          Submitted (Apr-2001) to the EMBL/GenBank/DDBJ databases.
CC            CC          -I- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = L-PYROLINE-5-
          CARBOXYLATE + NAD(P)H.
CC            CC          -I- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
CC            CC          -I- SIMILARITY: BELONGS TO THE PYROLINE-5-CARBOXYLATE REDUCTASE
          FAMILY.
CC            CC          -----
CC            CC          This SWISS-PROT entry is copyright. It is produced through a collaboration
CC            CC          between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC            CC          the European Bioinformatics Institute. There are no restrictions on its
CC            CC          use by non-profit institutions as long as its content is in no way
CC            CC          modified and this statement is not removed. Usage by and for commercial
CC            CC          entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC            CC          or send an email to licenselib@isb.sib.ch).
DR            DR          EMBL: Z77162; CAB00926.1; -.
DR            DR          EMBL: AE006952; AAA44743.1; -.
DR            DR          TIGR: MT0520; -.
DR            DR          TuberculoList; RV0500; -.
DR            DR          InterPro; IPR000304; P5CR.

```

```

0Y 48 LTRSNKTEHSHVJLEAKP- HIFPFLDEIGADV- -QARIWVSCAGVYTISSVE 107
      | : : : : : | | | | : : : : : | : : : : : | : : : : : |
Db 54 LYSADANVENAEFEVVAKKPADVEEPTADLANNTAAENDSEGEQFEVTVAGATIAFEE 113
      | : : : : : | | | | : : : : : | : : : : : | : : : : : |
0Y 102 KKLMAOPAPKVIKCHTNTPVVVOEGATVYATGHNALVERDGOLEQWMSVGFCEVEED 164
      | : : : : : | | | | : : : : : | : : : : : | : : : : : |
Db 114 SKLPACPTP- -VVRAMPNAALVYGAGVYALAKGFEVYTPOLEEVSALIDRAGCVLTPES 170
      | : : : : : | | | | : : : : : | : : : : : | : : : : : |
0Y 162 LIDAVTGLSGSPAPVAFAMLDH- -DGCKKKGCLPRRLATOLGAOALLGAKMILDSEOH- 217
      | : : : : : | | | | : : : : : | : : : : : | : : : : : |
Db 171 QLDNAVTVAGSGPAFFELLVALVDADGVGELSRQVATPLAQTAGASAMILLERMEDQ 230
      | : : : : : | | | | : : : : : | : : : : : | : : : : : |
0Y 218 -----PCOLDNWCSPGATIHAIHFLESGGFSRLINVAVASOIRREL 262
      | : : : : : | | | | : : : : : | : : : : : | : : : : : |
Db 231 GGANGELIMLQVLDLTPASRLRAAVTSPGCTTAALRLRENGEGRMAVDAVQAA----- 283
      | : : : : : | | | | : : : : : | : : : : : | : : : : : |
0Y 263 QSMADDEKSTP 273
      | : : : : : | | | | : : : : : | : : : : : | : : : : : |
Db 284 KSRSEQLRIITP 294
      | : : : : : | | | | : : : : : | : : : : : | : : : : : |

```

```

RESULT 15
PROC NEUR
ID PROC_NEUR STANDARD; PRT; 311 AA.
AC 012611
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
GN PRO-1.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96004701; PubMed=7565596;
RA Davis C.R., McPeck M.A., McClung C.R.;
RT "Molecular characterization of the proline-1 (pro-1) locus of
RT Neurospora crassa, which encodes delta 1-pyrroline-5-carboxylate
RT reductase."
RT Mol. Gen. Genet. 248:341-350(1995).
CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = 1-PYRROLINE-5-
CC CARBOXYLATE + NAD(P)H.
CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.tribe.ch/announce/).license@tribe.ch).
CC
DR EMBL: U030317; AAA83568.1; -.
DR InterPro: IPR000304; P5CR.
DR Pfam: PFD1089; P5CR; 1.
DR PROSITE: PS00521; P5CR; 1.
KW Oxidoreductase; Proline biosynthesis; NADP.
SQ SEQUENCE 311 AA; 32151 MW; 9A6D525D845984DF CRC64;

```

```

Qy 22 SAHIIISSEPMULPVYASLRKMGVNLRSRKETVYKSDVLFVAVKPHIIPETLDEIG-- 79
      | : : : : : | : | : : : : : | : : : : : | : : : : : |
Db 43 SVAKVEALSPLVKPSVSTLR---VLQSTSVNSAABADIIILGGCKPYMVGSLLSASGMK 99
      | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 80 -----ADVOARIHVYSCAGVYITISYVKKL-----NAFOPARKVYRCMTNPVYVO 127
      | : : : : : | : : : : : | : : : : : | : : : : : |
Db 100 DALTVKHTEGHARSOKIIISICAGTVYDLERVLREDVYGLSADNLPYVRAMPNTASKIR 155
      | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 126 EGATVYATGTHALVDE--COLLEQMLSSVGFCTEVEEDLIDATVGLSGSPAYAFMAID-- 187
      | : : : : : | : : : : : | : : : : : | : : : : : |
Db 160 ESMYTVINTYDPELPDVTYELLFTWIERGEVYVLLPRLIMACTSLCASSTAFALMMEBA 219
      | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 183 ADGGVYMGILPRRLAIQLGAAQLLGAANKLLDSEQHPCOLKDYNVCSFGGTTIHALCPLESG 244
      | : : : : : | : : : : : | : : : : : | : : : : : |
Db 220 ADGGVYMGILPRRAENARMAAQOTMRCGAAGIVLEGE--HPRALIREKYSTDGCGTIGGLVLYDEG 278
      | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 243 GFRSLIINVEASCI 257
      | : : : : : | : : : : : | : : : : : | : : : : : |
Db 279 GYAAVAARVAREATV 293

```

Search completed: January 31, 2002, 14:57:38  
Job time: 203 sec

Thu Jan 31 15:07:58 2002

us-09-912-717-1.rsp

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:53:30 ; Search time 16.15 Seconds

(without alignments)  
1481.041 Million cell updates/sec

Title: US-09-912-717-1

Perfect score: 1567  
Sequence: 1 MSVGFAGQLAYRFTAGT.....TPSSPGKLTLSLALGKKD 314

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database : PIR\_68:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1306	83.3	319	1 A41770	pyrroline-5-carbox
2	578	36.9	279	2 T23765	hypothetical prote
3	512.5	32.7	276	1 F02334	pyrroline-5-carbox
4	501.5	32.0	274	1 S10186	pyrroline-5-carbox
5	479	30.6	273	2 T06477	probable pyrroline
6	423	27.0	284	2 T36286	pyrroline-5-carbox
7	421	26.9	267	2 S76767	probable pyrroline
8	406.5	25.9	269	2 H85533	pyrroline-5-carbox
9	402.5	25.7	269	1 RDECC3	pyrroline-5-carbox
10	377	24.1	273	2 J00418	pyrroline-5-carbox
11	345	22.0	270	2 E69682	pyrroline-5-carbox
12	326.5	20.8	294	2 S72897	pyrroline-5-carbox
13	324.5	20.7	278	1 G69964	pyrroline-5-carbox
14	323.5	20.6	261	2 JC2078	pyrroline-5-carbox
15	323.5	20.6	279	2 G83837	pyrroline-5-carbox
16	319	20.4	255	2 E72360	pyrroline-5-carbox
17	306.5	19.6	295	2 G70745	probable proc prot
18	301.5	19.2	311	2 S57863	pyrroline-5-carbox
19	297.5	19.0	305	2 C82524	pyrroline-5-carbox
20	297	19.0	265	2 F70315	pyrroline carboxyl
21	296	18.9	320	2 JC4830	pyrroline-5-carbox
22	295.5	18.9	262	2 D86860	pyrroline-5-carbox
23	293	18.7	266	2 G72769	Delta 1-pyrroline-
24	292.5	18.7	282	2 T50305	pyrroline-5-carbox
25	289	18.4	264	2 C75385	pyrroline-5-carbox
26	281.5	18.0	286	2 S25293	probable pyrroline
27	279	17.8	263	2 D71281	hypothetical prote
28	266	17.0	299	2 T29226	pyrroline-5-carbox
29	264	16.8	271	2 I64060	pyrroline-5-carbox

30	258.5	16.5	272	2 D82321	pyrroline-5-carbox
31	229	14.6	267	2 G83760	pyrroline-5-carbox
32	212.5	13.6	263	2 H82015	probable pyrroline
33	212.5	13.6	263	2 E81243	pyrroline-5-carbox
34	193.5	12.3	257	2 F64664	pyrroline-5-carbox
35	192.5	12.3	243	2 G81310	probable pyrroline
36	179.5	11.5	257	2 E71850	probable pyrroline
37	179	11.4	275	2 G83816	non-essential gene
38	162.5	10.4	273	2 F69602	pyrroline-5-carbox
39	148.5	9.5	253	2 H69219	pyrroline-5-carbox
40	114.5	7.3	264	1 F69100	conserved hypothet
41	106.5	6.8	1043	2 H83329	probable RND efflu
42	106	6.8	348	2 A32512	glycerol-1-3-phosfa
43	104	6.6	272	1 F69855	pyrroline-5-carbox
44	104	6.6	451	2 S26441	hypothetical prote
45	102	6.5	349	2 A25952	glycerol-3-phospha

#### ALIGNMENTS

RESULT 1

A41770 pyrroline-5-carboxylate reductase (EC 1.5.1.2) - human

N:Alternate names: P5C reductase  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: A41770  
R:Dougherty, K.M.; Brandtiss, M.C.; Valle, D.  
J. Biol. Chem. 267, 871-875, 1992

A:Title: Cloning human pyrroline-5-carboxylate reductase cDNA by complementation in S

A:Reference number: A41770; M01D:92112821  
A:Accession: A41770  
A:Molecule type: mRNA

A:Residues: 1-319 <DOU>  
A:Cross-references: GB:M77836; NID:q189497; PIDN:AAA36407.1; PID:q189498  
A>Note: sequence extracted from NCBI Backbone (NCBIN:75606, NCBIIP:75608)

A:Gene: GDB:PYCRL; P5C; PYCR  
A:Cross-references: GDB:135716; OMIM:179035  
A:Map position: 17pter-17qter  
C:Superfamily: pyrroline-5-carboxylate reductase

C:Keywords: oxidoreductase; proline biosynthesis

Query Match 83.3%; Score 1306; DB 1; Length 319;  
Best Local Similarity 82.7%; Pred. No. 2.1e-96;  
Matches 263; Conservative 27; Mismatches 22; Indels 6; Gaps 2;

QY	1	MSVGFAGQLAYR---FTAGILSAHKITASSPENMLPTVSALRMKGYNLTPRSNKETV	56
DB	1	MSVGFAGQLAFALAKGFTAGVLAHKIMASSPMDLAVSALRMKGYNLTPHNKETV	60
QY	57	KHSVDYFLAVRPHIIFLDEIGADVQARHIVSCAGVTTSSVEKKLMAFOPAPKPIRC	116
DB	61	QHSVDYFLAVRPHIIFLDEIGADVQARHIVSCAGVTTSSVEKKLMAFOPAPKPIRC	120
QY	117	MTNPVVVQEGATVYATGTHALVEDGOLTEOLMSSVGCFTVEEDLLDAVYGLSGSPAY	176
DB	121	MTNPVVVQEGATVYATGTHALVEDGOLTEOLMSSVGCFTVEEDLLDAVYGLSGSPAY	180
QY	177	AFMALD--ADGVKNGLPRLRIOLGAOLIGAAMKMLDSQHRQCOLKDNVCSFGATIH	234
DB	181	AFMALDADALADGVKNGLPRLRIOLGAOLIGAAMKMLDSQHRQCOLKDNVCSFGATIH	240
QY	235	ALHPLSESGFRSLINAVEASCIRTRLOSMADDEKISPAALKTLLDRKLESPYVSTL	294
DB	241	ALHPLSESGFRSLINAVEASCIRTRLOSMADDEKISPAALKTLLDRKLESPYVSTL	300
QY	295	TPSSPGKLTLSLALGK 312	
DB	301	SPSGHTKLTLSLALGK 318	



Db 247 IHELENGCFRGITLMNVAAAKRSREL 273

RESULT 5

T06477

probable pyrroline-5-carboxylate reductase (EC 1.5.1.2) - garden pea

C:Species: Pisum sativum (garden pea)

C:Date: 23-Apr-1999

#sequence\_revision 23-Apr-1999

#text\_change 11-Jun-1999

C:Accession: T06477

R:Williamson, C.L.; Slocum, R.D.

Plant Physiol. 100, 1464-1470, 1992

A:Title: Molecular cloning and evidence for osmoregulation of the di-pyrroline-5-carboxy

A:Reference number: 215706

A:Accession: T06477

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-273 <M12>

A:Cross-references: EMBL:X62842; NID:g20850; PID:CAAA4646.1; PID:g20851

A:Experimental source: cv. Mando

C:Genetics:

A:Gene: PROC

C:Superfamily: pyrroline-5-carboxylate reductase

C:Keywords: oxidoreductase; proline biosynthesis

Query Match

Best Local Similarity 40.1%; Pred. No. 9.3e-31;

Matches 108; Conservative 55; Mismatches 90; Indels 16; Gaps 5;

Db 12

SVGFIGAGGLAYRFR---TAAGILSAHKITA--SPEKMLPTVSAIRKMGVNLTSNKT 55

12 TIGFIDAGMMAISIAKAGARSGLVSSRIVTASHNSRR---AAFEISGITVLSNDV 67

56 VKHSDVLEFLAVKPHIIPILDEIGADVQARHIVSCAAGVTISVEKKIMAFQAPKVR 115

68 VRASNVVSVKPOLKVDVLTAKLPLTKDKLLVSAAGIKL---KDLQENAGHERETR 123

116 CMTNPVVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEDLIDAVTGLSGSPA 175

124 VPMNPAAVGAQAAVWSLGCATTEEDANLISOLFGSIKIMKADKFPDALITGLSGSPA 183

176 YAFMALD--ADGVKMGIPRRRLAIOGAQALLGAAMKMLDSEQRHCOLKDNVCSFGAT 233

184 YITALEALADGVAAAGLPRLDALSLASQTVLGAASMATLSGRHGOGLKDDVTSFGTTI 243

234 HALHFLSGGFRSLILINAVEASCIRTEL 262

244 AGVHELEKCGFRGTLTMNVAAAKRSREL 272

RESULT 6

T36286

pyrroline-5-carboxylate reductase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999

#sequence\_revision 03-Dec-1999

#text\_change 21-Jan-2000

C:Accession: T36286

R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1998

A:Reference number: 221603

A:Accession: T36286

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-284 <SE8>

A:Cross-references: EMBL:AL049819; PIDN:CA842663.1; GSPDB:GND0070; SCOEDB:SCET.04C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: PROC; SCOEDB:SCET.04C

C:Superfamily: pyrroline-5-carboxylate reductase

Query Match

Best Local Similarity 40.1%; Pred. No. 2.8e-26;

Score 423; DB 2; Length 284;

Matches 108; Conservative 38; Mismatches 89; Indels 34; Gaps 5;

Db 4

GFIGAGGLAYRFR---TAAGILSAHKITA--SPEKMLPTVSAIRKMGVNLTR-----SNKT 55

34 GMIGAG-----WAPADL-----VTARRRERADELARHRCVTPVTNAA 72

56 VKHSDVLEFLAVKPHIIPILDEIGADVQARHIVSCAAGVTISVEKKIMAFQAPKVR 115

73 AKADTLLITVKKPQDMGTLDELAPHPADRLVSGAAGVPTSFEEERL---APGTPVR 129

116 CMTNPVVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEDLIDAVTGLSGSPA 175

130 VMTNPALVDAMSVISAGTATTAHLTHEIFGAAGKTRVPESSQDQACTALSGSPA 189

176 YAFMALD--ADGVKMGIPRRRLAIOGAQALLGAAMKMLDSEQRHCOLKDNVCSFGAT 233

190 YEFYLVEMTDAIGLILGPRDKAHDLVQSAIGAAKMLRSGEHPVRLRENVSPACTTI 249

234 HALHFLSGGFRSLILINAVEASCIRTEL 262

250 NAIRELEHNGVRAALIALEAARDRSREL 278

RESULT 7

S76767

probable pyrroline-5-carboxylate reductase (EC 1.5.1.2) - Synecocystis sp. (strain P

C:Species: Synecocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997

#sequence\_revision 25-Apr-1997

#text\_change 20-Jun-2000

A:Accession: S76767

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yae

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis

A:Reference number: S74322; MUID:97061201

A:Accession: S76767

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-267 <KAN>

A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BA18679.1; PID:g165

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Start codon: GTG

C:Superfamily: pyrroline-5-carboxylate reductase

C:Keywords: oxidoreductase; proline biosynthesis

Query Match

Best Local Similarity 38.5%; Pred. No. 3.7e-26;

Matches 106; Conservative 47; Mismatches 96; Indels 26; Gaps 8;

Db 1

MSVFIGAGGLA---YRTPAAGILSAHKITA--SPEKMLPTVSAIRKMGVNLTR-----SNKT 51

3 IQDLITGGVAAEALILRLAIEKTAPEEITVGPB-----GAKRRTYLOKTYQVARS 56

52 NKEIVKHSVLEFLAVKPHIIPILDEIGADVQARHIVSCAAGVTISVEKKIMAFQAPKVR 111

57 NOEAAVSEVLLAVKPOVDRLVSLAGAN--RPLVISILAGVISIORIGKGF-----PDH 111

112 KVIRCMNTNPVVQEGATVYATGTHALVEDGQL--LEQLMSSVGFCTEVEDLIDAVTGL 169

112 ATRAMPPTPATV--GAGMTAIAANKKVEEDPQLAKAAVIFSAVGNVVEENIMDAVTGV 169

170 SGSGPAVAFMALD--ADGVKMGIPRRRLAIOGAQALLGAAMKMLDSEQRHCOLKDNVCS 227

170 SGSGPAVVALMIEALADGCVLAGLPRALAKLQVTLGNAELIKETREHNPADIKKQVTS 229

228 PGCATTHALHFLSGGFRSLILINAVEASCIRTEL 262

230 PGCTTIAGVAVLEKMGFRSAIIEAVRAAYRSOEL 264





Db	19	VAEIGAGSMA - GCGMTSGIVIRANKIPKONITCVTRNSMTERTELEDELOYGIGALPNQJCIE	77
QY	58	HSVULAKPPIHPIFIDEIGADVQARHIVYSCAGVITSSVEKKIAPAPAKVIRCM	117
Db	78	DMVULILAKMPDAENALSSLSKSRIOHQHILSVLAGITTSIESQLNEQP---VVRVM	134
QY	118	TNFPYVVGAGAVYVATGTHALVEQQLLEDMSSVFCVEVEDLIDAVTGLSGSGPAYA	177
Db	135	PNITSMIGASNAIILGKTVYSESDIKTAELALGCMGEVYTTIDENMDIFPTIASSGPAYP	194
QY	178	FMALD--ADGVKMGMLPRLAIQIAGCALIGAAMKMLDSEHQPCOLKDNVCSPGATITHA	235
Db	195	YVMEITEKTGEAGLIDQLRSIGAOULTLGAAMKMLMETGEHPEILRDNITSPNGTTAAG	254
QY	236	LHFL-ESGGFRSL 248	
Db	255	LQALKKSGGKRFLL 268	
RESULT	12		
	772897		

[illegible]

R.Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beret, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 300, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier, A.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.F.; Koester, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidos, A.; Lardinois, Y.; M. Ogawa, K.; Ogiwara, A.; Oudega, B.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sekowski, A.; Seton, A.; T. Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yano, K.; Yata, K.; Yoshida, K.; A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Yanchin, A.

A:Title: The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. A:Reference number: A69580; MUID:98044033

A:Accession: G69964

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-278 <KUN>

A:Cross-references: GB:299116; GB:AL009126; NID:92634723; PIDN:CA14312.1; PID:92634815

A:Experimental source: strain 168

C:Genetics:

A:Gene: ynfj0

C:Superfamily: pyrroline-5-carboxylate reductase

C:Keywords: oxidoreductase

Query Match 20.7%; Score 324.5; DB 1; Length 278;  
Best Local Similarity 29.9%; Pred. No. 1.8e-18;  
Matches 88; Conservative 54; Mismatches 113; Indels 39; Gaps 7;

QY 3 VGTGAGQLAYR---FTAGCISAKKIASSPBNMLPVSAIRKMGVLTSSNKEVYKH 58  
DB 4 IGFAGASMAEMINGILGICRPHIYITNSNDERIETKEVSPCKDKKEFTFH 63  
QY 59 SDVFLAVKP-----HIIFLDEICADVQARIIVSCAGVTTSSVE---KKIMA 106  
DB 64 TDTITLAFKRDAAESIDSTRPIKQ-----LVISVLAGLITETIHYFGKRL-- 112  
QY 107 FQPAKYIRCMNTPVVVOGATVATGTHALVEDQQLLEQMLSSVGFCTEVEDLIDAV 166  
DB 113 -----AVIRVPMNTSAIRKSATGFSVSTASKNDIIMAKALLETIGDALTVEERLIDAV 167  
QY 167 TGLSGSPAYAFMLDA--DGVYKMGPRRLAIQIGAOALLGAKMILDSQHPQCKDN 224  
DB 168 TAIAGSGPAYRYTEAMEKRAQKVGDKETAKALIIQTMAGTIDMLDSGKPELKRK 227  
QY 225 VCSPEGATIALHFLFESGFRSLINAVEASCIRTRRELQSMADQEKISPAALK 278  
DB 228 ITSPGCTEAGLRALDLSRFEALIHCIETAKRSAEI-----KEOFAGAMLER 276

RESULT 14  
JC2078  
pyrroline-5-carboxylate reductase (EC 1.5.1.2) - *Thermus aquaticus* (strain HB27)  
C:Species: *Thermus aquaticus*  
C:Date: 30-Sep-1993 #sequence\_revision 20-Aug-1994 #text\_change 26-Aug-1999  
C:Accession: JC2078  
R.Hoshino, T.; Kosuge, T.; Hidaka, Y.; Tabata, K.; Nakahara, T.  
Biochem. Biophys. Res. Commun. 199, 410-417, 1994  
A:Title: Molecular cloning and sequence analysis of the *proC* gene encoding delta-1-pyrroline-5-carboxylate reductase  
A:Reference number: JC2078; MUID:94168609  
A:Accession: JC2078  
A:Molecule type: DNA  
A:Residues: 1-261 <HOS>  
C:Superfamily: pyrroline-5-carboxylate reductase  
C:Keywords: oxidoreductase

Query Match 20.6%; Score 323.5; DB 2; Length 261;  
Best Local Similarity 33.9%; Pred. No. 2e-18;  
Matches 93; Conservative 45; Mismatches 107; Indels 29; Gaps 11;

QY 1 MSVGTGAGQLAYRPTA---GILSANK--IIASSPBNMLPVSAIRKMGVLTSSNKEVYKH 53  
DB 1 MLTAVVAGLAKMRSLTKALRGRFLRPEVGLRTERPSR---ELAEPPGVPLTRAD- 56  
QY 54 ETYKSHDVLAVKPHIIFLDEIGADYQARHIV--VSCAGVTTSSVKKIMARPOPAV 111  
DB 57 --LGMARVLLAVQRPDPALAPET---AHNRIGYISIMAGISTVSLARRL-----DNR 105  
QY 112 KYIRCMNTPVVVOGATVATGTHAL-VEGQQLLEQMLSSVGFCTEVEDLIDAVTGLS 170  
DB 106 RYVRAAPNLAIVIGSSTALTAKEAREADLAFATVGDVYEIPEHLPDFTGMS 165  
QY 171 GSGPAYAFMLD--ADGVYKMGPRRLAIQIGAOALLGAKMILDSQHPQCKDNVCS 228  
DB 166 ASAPALAVVAELADVAGYKMGPRRLALRLADALATGELT--KGRHPQVQDEVAS 223

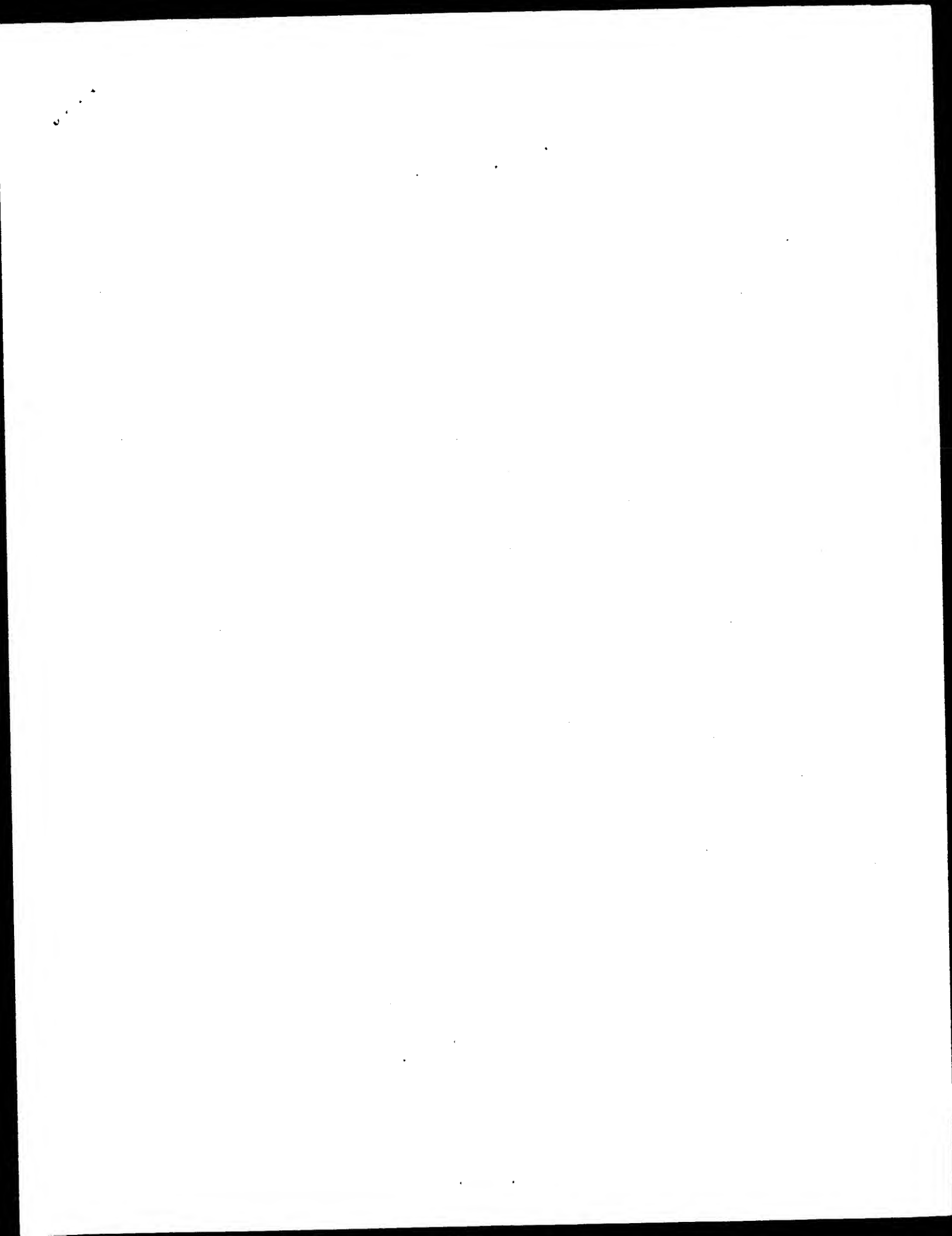
RESULT 15  
G83837  
pyrroline-5-carboxylate reductase BH1503 [imported] - *Bacillus halodurans* (strain C-1)  
C:Species: *Bacillus halodurans*  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: G83837  
R.Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H. Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a  
A:Reference number: A83650; MUID:20263314  
A:Accession: G83837  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-279 <STO>  
A:Cross-references: GB:AP001512; GB:BA000004; NID:910174030; PIDN:BA05222.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH1503  
C:Superfamily: pyrroline-5-carboxylate reductase

Query Match 20.6%; Score 323.5; DB 2; Length 279;  
Best Local Similarity 33.2%; Pred. No. 2.2e-18;  
Matches 91; Conservative 47; Mismatches 125; Indels 11; Gaps 5;

QY 3 VGTGAGQLAYRPTAAGILSANKIASSP-EMNLPVSAIRKMGVLTSSNKEVYKH 57  
DB 7 ITFAGASMA-ESTIAGLCKILLPSOVATAMICDEAKITLEDRYGICTONROEAVR 65  
QY 58 HSDVFLAVKPHIIFLDEIGADYQARHIVSCAGVTTSSVEKKIMARPOPAVIRCM 117  
DB 66 QGTITFLAMKPKNIEAEIEIRGETEIKOLFISVLAGITTSYETLLAEV---VVRIM 122  
QY 118 TNPVVOGATVATGTHALVEDQQLLEQMLSSVGFCTEVEDLIDAVTGLSGSPAYA 177  
DB 123 PNTSKAYASATGICGGRYANASHVKIATLIFATIGVTEVEDKIDAVTGLAGSPAYF 182  
QY 178 FMAIDA--DGVYKMGPRRLAIQIGAOALLGAKMILDSQHPQCKDNVCSGPGATIA 235  
DB 183 YVWEAMRAVAVOSGLTESEKAFISQTLIGTGRRREQISYTAEDLYKEVMSFGITENG 242  
QY 236 LHFLESGFRSLINAVEASCIRTRRELQSMADQ 269  
DB 243 LNVLEQDQMOAIEEATITAAINRSRELGSTPSLE 276

Search completed: January 31, 2002, 14:54:14  
Job time: 44 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:53:55 ; Search time 24.93 seconds

(without alignments)  
1842.339 Million cell updates/sec

Title: US-09-912-717-1

Sequence: 1 MSVGFAGAGLAFRTAFTACT.....TPSSPKLTLRLSLALGKKMD 314

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: SPREMBL\_17:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhnc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1231.5	78.6	319	4	O9Y5J4 homo sapien
2	898	57.3	272	4	O9H8Q4
3	578	36.9	279	5	O21544
4	537.5	34.3	274	4	O9H896
5	522.5	33.3	274	11	O9DCC4
6	503	32.1	273	5	O9VEJ3
7	500.5	31.9	274	10	O9AVM3
8	481	30.7	280	5	O9V3F8
9	469	29.9	280	5	O96643
10	456	29.1	270	1	O9H899
11	448	28.6	266	2	O87725
12	433	27.0	284	2	O9X8G1
13	399	25.5	272	2	O9K308
14	380	24.3	248	11	O9DOX2
15	376	24.0	254	5	O9NGS4
16	344	22.0	279	1	O9H896
17	325.5	20.8	273	1	O9H896
18	323.5	20.6	279	2	O9KCR6
19	322	20.5	209	5	O27722

20	319	20.4	255	2	O9W247
21	301	19.2	290	2	O9E283
22	297.5	19.0	305	2	O9PA08
23	295.5	18.9	262	2	O9CEP7
24	293	18.7	266	1	O9YFV4
25	292.5	18.7	282	3	O9P777
26	289	18.5	275	2	O9CPE8
27	289	18.4	264	2	O9RUT7
28	263.5	16.8	271	2	O9FUD7
29	258.5	16.5	260	2	O9AAU9
30	258.5	16.5	272	2	O9KU05
31	258	16.5	256	2	O9A1S9
32	252.5	16.1	271	2	O9P720
33	229	14.6	267	2	O9KEG5
34	212.5	13.6	263	2	O9K1N1
35	212.5	13.6	263	2	O9JWM2
36	192.5	12.3	243	2	O9PML8
37	179	11.4	275	2	O9KDB6
38	151.5	9.7	117	10	O41907
39	151	9.6	99	2	O07508
40	148.5	9.5	253	1	O26983
41	114.5	7.3	264	1	O27779
42	106.5	6.8	250	4	O95435
43	106.5	6.8	1043	2	O910V6
44	101.5	6.5	350	13	O98S39
45	101	6.4	8817	2	O53840

## ALIGNMENTS

RESULT 1  
ID O9Y5J4 PRELIMINARY: PRT: 319 AA.  
AC O9Y5J4  
DT 01-NOV-1999 (TREMURel. 12, Created)  
DT 01-NOV-1999 (TREMURel. 12, Last sequence update)  
DT 01-JUN-2001 (TREMURel. 17, Last annotation update)  
DE PYROLINE 5-CARBOXYLATE REDUCTASE ISOFORM.  
GN P5CR2.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA KOK L.D.S., Au T.C.C., Fung K.P., Tsui S.K.W., Lee C.Y., Waye M.M.Y.;  
RT "Molecular characterization of a novel human pyroline-5-carboxylate  
reductase isolated in liver."  
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF151351; AAD34611.1; -  
DR InterPro: IPR000304; P5CR.  
DR Pfam: PF01089; P5CR; 1.  
DR PROSITE: PS00521; P5CR; 1.  
SQ SEQUENCE 319 AA: 33726 MW: 01D497002A598838 CRC64;

Query Match 78.6% Score 1231.5; DB 4; Length 319;  
Best Local Similarity 81.7% Pred. No. 3.8e-90;

Matches 263; Conservative 7; Mismatches 41; Indels 11; Gaps 4;

OY 1 MSVGFAGAGLAFRTAFTACT...IASSPEMNLPTVSLRKGVNLTRSNKEFKHS 59  
DB 1 MSVGFAGAGLAFRTAFTACT...SRPQIPVGSQDQNSQLPRNPAHGVRAQEDCEPDTQOG 57  
OY 60 D-----VLEIAVKPHLIPILDEIGADYQARHIVSCAGVTISSVEKKLMAFOPAPKVI 114  
DB 58 DDEAQRPRPVSGEAIHIFILVEIGADYQARHIVSCAGVTISSVEKKLMAFOPAPKVI 117  
OY 115 RQMTNPVVOVGATVATGTHALVEDGQLLEQLMSSVGFCTVEEDLITAVGLSGSP 174  
DB 118 RQMTNPVVOVGATVATGTHALVEDGQLLEQLMSSVGFCTVEEDLITAVGLSGSKP 177

QY 175 AYAFMAMD--ADGCVKMGUPRRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 232  
 DB 178 AYAFMAMDALADGCVKMGUPRRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 237  
 QY 233 IHALHFLSEGGFRSLINAVASCIRTRRELOSMADQEKISPAALAKTLLDRVYKLESPYVS 292  
 DB 238 IHALHFLSEGGFRSLINAVASCIRTRRELOSMADQEKISPAALAKTLLDRVYKLESPYVS 297  
 QY 293 TLTPSSPGKLLTRSLALGCKKD 314  
 DB 298 TLTPSSPGKLLTRSLALGCKKD 319  
 RESULT 2  
 Q9HB04 PRELIMINARY; PRT: 272 AA.  
 AC Q9HB04;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE HYPOTHETICAL 28.5 KDA PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gu J.R., Man D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,  
 RA Gu W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,  
 RA Yu J., Han L.H.;  
 RT "Novel Human cDNA clones with function of inhibiting cancer cell  
 growth";  
 RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AF218000; AAG17242.1; -  
 DR InterPro: IPR000304; P5CR.  
 DR Pfam: PF01089; P5CR; 1.  
 DR Hypothetical protein.  
 KW SQUENDE 272 AA; 28471 MW; 0BC71EA751F78F4 CRC64;  
 SO  
 Query Match 57.3%; Score 898; DB 4; Length 272;  
 Best Local Similarity 83.7%; Pred. No. 1e-63;  
 Matches 180; Conservative 19; Mismatches 10; Indels 6; Gaps 2;  
 QY 1 MSVGFAGAGOLAYR---FTAGILSAHKTIASSPEMNLPTYSALPKMGVNLTRSKETV 56  
 DB 1 MSVGFAGAGOLAYR---FTAGILSAHKTIASSPEMNLPTYSALPKMGVNLTRSKETV 60  
 QY 57 KHSVDFLAVKPHIIPFLIDEIGADVOARHIVYSCAAGVITSSVEKKLMAFOAPKVIK 116  
 DB 61 QHSVDFLAVKPHIIPFLIDEIGADVOARHIVYSCAAGVITSSVEKKLMAFOAPKVIK 120  
 QY 117 MTNTPVVOGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSPAY 176  
 DB 121 MTNTPVVOGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSPAY 180  
 QY 177 AFMAMD--ADGCVKMGUPRRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 209  
 DB 181 AFMAMD--ADGCVKMGUPRRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 215  
 RESULT 3  
 Q21544 PRELIMINARY; PRT: 279 AA.  
 AC Q21544;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE M153x1 PROTEIN.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Matthews P.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kersey J., Kistner J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy A., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smailson N., Smith A., Sonhammer E., Staden K., Sulston J.,  
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Spratt J., Woldman P.;  
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL: Z67995; CAB91943.1; -  
 DR InterPro: IPR000304; P5CR.  
 DR Pfam: PF01089; P5CR; 1.  
 SO SEQUENCE 279 AA; 29165 MW; C68CAD4BC4EB20E6 CRC64;  
 Query Match 36.9%; Score 578; DB 5; Length 279;  
 Best Local Similarity 44.6%; Pred. No. 3.2e-38;  
 Matches 127; Conservative 55; Mismatches 89; Indels 14; Gaps 4;  
 QY 1 MSVGFAGAGOLAYR---FTAGILSAHKTIASSPEMNLPTYSALPKMGVNLTRSKETV 56  
 DB 1 MSVGFAGAGOLAYR---FTAGILSAHKTIASSPEMNLPTYSALPKMGVNLTRSKETV 60  
 QY 57 KHSVDFLAVKPHIIPFLIDEIGADVOARHIVYSCAAGVITSSVEKKLMAFOAPKVIK 116  
 DB 61 QHSVDFLAVKPHIIPFLIDEIGADVOARHIVYSCAAGVITSSVEKKLMAFOAPKVIK 117  
 QY 117 MTNTPVVOGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSPAY 176  
 DB 118 MTNTPVVOGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSPAY 177  
 QY 177 AFMAMD--ADGCVKMGUPRRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 234  
 DB 178 AFMAMD--ADGCVKMGUPRRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 237  
 QY 235 ALHFLSEGGFRSLINAVASCIRTRRELOSMADQEKISPAALAKT 279  
 DB 238 ALHFLSEGGFRSLINAVASCIRTRRELOSMADQEKISPAALAKT 277  
 RESULT 4  
 Q9H896 PRELIMINARY; PRT: 274 AA.  
 AC Q9H896;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE CDNA FLJ13852 FTS, CLONE THYR0100934, WEAKLY SIMILAR TO PYRROLINE-5-  
 DE CARBOXYLATE REDUCTASE (EC 1.5.1.2).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=THYROID GLAND;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,  
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,  
 RA Arita M., Nadekura T., Ishii S., Kawai Y., Saito Y., Yamamoto J.,  
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y., Oshima A.;

"MEDD human cDNA sequencing project."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK023914; BAB14721.1; -  
 DR InterPro: IPR000304; P5CR.  
 DR Pfam: PF01089; P5CR.1.  
 DR PROSITE: PS00521; P5CR.1.  
 DR SEQUENCE 274 AA; 28663 MW; 846DE6C03F3B548 CRC64;

Query Match 34.3%; Score 537.5; DB 4; Length 274;  
 Best Local Similarity 45.3%; Pred. No. 5; Le-35;  
 Matches 121; Conservative 45; Mismatches 88; Indels 13; Gaps 5;

QY 3 VGFAGAGCQV---LAFRTAGILSAHKIIASSP-EMNLPTVSAIRKMGVLTIRSNKTEVK 57  
 Db 11 VGFVAGAGMGALAGLIRACKVAGAHILASAPTRDLNCRHAL---SCRTTHSNQVLDQ 67  
 QY 58 HSDVFLAVKPHIIPFLIDEGADVQARHIVSCAGVTTSSVEKKIMAFQAPKVIKCM 117  
 Db 68 SCLVIFATKPHVLPVLAELAVAPVTTEHILVSAAGVSLTEBEL---PPNTRVLRLV 124  
 QY 118 TNPVVOEGATVATGTHALVEDGQLLEQMSVGCETEVEEDLDAVGTLSGSPAY- 176  
 Db 125 PNLPCVVOEGAVIMARKHGVSSFTKLLQHLLEACGCEEPVAVYDIHTGLSGGVAFCV 184  
 QY 177 AFMALDADGCVKMGKLPRLAIOGAQALLGAARKMLDSEQHPCQLKDNVCSFGATTHA 235  
 Db 185 CAFSALALAEAGAVKMGKMPSSLAHRIAAQTLTGCTAKMLHEGHPAQDLSDVCTPGTITYG 244  
 QY 236 LHFLESGFSLINAVEASCITRREL 262  
 Db 245 LHALEGGGLRAATWSAVEAATCRAKEL 271

RESULT 5  
 ID 09DCC4 PRELIMINARY; PRT; 274 AA.  
 AC 09DCC4;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE 1110058B13RIK PROTEIN.  
 GN 1110058B13RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=KIDNEY;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinnagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 Arkawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
 Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 Blake J., Botfield D., Boujuna N., Carninci P., de Bonaldo M.F.,  
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 Guernstein S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,  
 Lyons P., Marchionni L., Mashima J., Mazzei J., Mombert P.,  
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 Saeki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 Rayshav-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,  
 Hayashizaki Y.;  
 RA \*Functional annotation of a full-length mouse cDNA collection.\*;  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK002912; BAB2451.1; -  
 DR MGI:1913444; 1110058B13RIK.

DR InterPro: IPR000304; P5CR.  
 DR Pfam: PF01089; P5CR.1.  
 DR SEQUENCE 274 AA; 28694 MW; FE3892C01C6068A5 CRC64;

Query Match 33.3%; Score 522.5; DB 11; Length 274;  
 Best Local Similarity 42.9%; Pred. No. 8e-34;  
 Matches 114; Conservative 47; Mismatches 94; Indels 11; Gaps 4;

QY 3 VGFAGAGCQV---LAFRTAGILSAHKIIASSP-EMNLPTVSAIRKMGVLTIRSNKTEVK 58  
 Db 11 VGFVAGAGMGALAGLIRACKVAGAHILASAPTRDLNCRHAL---SCRTTHSNQVLDQ 68  
 QY 59 SDVFLAVKPHIIPFLIDEGADVQARHIVSCAGVTTSSVEKKIMAFQAPKVIKCM 118  
 Db 69 SCLVIFATKPHVLPVLAELAVAPVTTEHILVSAAGVSLTEBEL---PPNTRVLRLV 125  
 QY 119 NTPVVOEGATVATGTHALVEDGQLLEQMSVGCETEVEEDLDAVGTLSGSPAY- 176  
 Db 126 PNLPCVVOEGAVIMARKHGVSSFTKLLQHLLEACGCEEPVAVYDIHTGLSGGVAFCV 185  
 QY 177 AFMALDADGCVKMGKLPRLAIOGAQALLGAARKMLDSEQHPCQLKDNVCSFGATTHA 236  
 Db 186 TFSALALAEAGAVKMGKMPSSLAHRIAAQTLTGCTAKMLHEGHPAQDLSDVCTPGTITYG 245  
 QY 237 LHFLESGFSLINAVEASCITRREL 262  
 Db 246 LHALEGGGLRAATWSAVEAATCRAKEL 271

RESULT 6  
 ID 09VEJ3 PRELIMINARY; PRT; 273 AA.  
 AC 09VEJ3;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE CG5840 PROTEIN.  
 GN CG5840.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Pnytoidea; Dirosophillidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams W.D., Celniker S.E., Holt R.A., Evans C.A., Iyer J.D.,  
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins J., Galie R.F.,  
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 Anil J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,  
 Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borkova D., Borkan M.R., Bouck J., Brokstein P., Brotler P.,  
 Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,  
 Juelali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

Query Match	31.98;	Score 500.5;	DB 10;	Length 274;
Best Local Similarity	41.58;	Pred. No. 4.5e-32;		
Matches 112;	Conservative 51;	Mismatches 90;	Indels 17;	Gaps 5

RA Styrskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Massarman D.A., Weinsbach G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.



RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Glibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN 121

RP SEQUENCE FROM N.A.

RC STRAIN-CS:

RA Misenar S.R., Walker V.K.;

RT "High density of unrelated genes showing overlapping and intratronic  
 RT transcription units in *Drosophila*.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: A6003724; AAF55626.1; -  
 DR EMBL: AF170829; AAD49740.1; -  
 DR FlyBase: FBgn0015781; P5CR.

DR InterPro: IPR000304; P5CR.

DR Pfam: PF01089; P5CR; 1.

DR PROSITE: PS00521; P5CR; 1.

KW Oxidoreductase.

SO SEQUENCE 280 AA; 29621 MW; D3D4PD9611574684 CMC64;

Query Match 30.7%; Score 481; DB 5; Length 280;  
 Best Local Similarity 39.1%; Pred. No. 1.7e-30;

Matches 108; Conservative 51; Mismatches 95; Indels 22; Gaps 6;

QY 3 VGFIGAGQLAYR---FTAAGILSAHKIIASSPEKMLPTVSALKRKGYNLTRSKNETVKH 58  
 DB 8 IGFIGGNMVAIGSLVRCGIVKASQVQVSP--HIENLQMRDLCAVTCDDCMVLEH 65  
 QY 59 SDVFLAVKPHIIPILDEIGADVQARHI-----VSCAAGVTISVEKKLMFOP 109  
 DB 66 SDIVFCVAPR---MLTPCAALQKVKHVPASAKDASKLVSVLADTSLTELEAFSPFGS 121  
 QY 110 AP-KVIRCMTNTPVVVGEGATVATGTHALVEDGQLLEQLMSSVGFCTEVEDLIDAVTG 168  
 DB 122 SELKVIPTMPTSMQVGECTVYTGNAVSHHDEKITHLNLGLAQVPESMIDAVTG 181  
 QY 169 LSGSPATVAFMAD--ADGVKMGLPRLAIOGAQALLGAAMKMLDSEQHPQQLKDNVC 226  
 DB 182 VAGCGPAFYVYITIEALADGVKQGVPRQMALQFAOITLGAATVLLTGKHPAVLREVC 241  
 QY 227 SPGATIHAIHFLESGFRSLINAVEASCIRREL 262  
 DB 242 SPGATIVGVHELEKGNLRSTLINAVEKSSORSABL 277

RESULT 9

ID 096643 PRELIMINARY: PRT; 280 AA.

AC 096643;

DT 01-MAY-1999 (Tremblrel. 10, Created)

DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)

DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)

OS PYRROLINE 5-CARBOXYLATE REDUCTASE.

GN *Drosophila melanogaster* (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Ephygrotidae; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC NCBI\_Taxid=7227;

RN 111

RP SEQUENCE FROM N.A.

RC STRAIN-CS:

RA Misenar S.R., Walker V.K.;

RT "Cloning and sequence analysis of the gene encoding pyrroline 5-  
 RT carboxylate reductase from *Drosophila melanogaster*.";  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF098020; AAC70780.1; -  
 DR FlyBase: FBgn0015781; P5CR.  
 DR InterPro: IPR000304; P5CR.  
 DR Pfam: PF01089; P5CR; 1.  
 DR PROSITE: PS00521; P5CR; 1.  
 SO SEQUENCE 280 AA; 29688 MW; 2A7416D39D2A2369 CRC64;

Query Match 29.9%; Score 469; DB 5; Length 280;  
 Best Local Similarity 38.4%; Pred. No. 1.5e-29;  
 Matches 106; Conservative 51; Mismatches 97; Indels 22; Gaps 6;

QY 3 VGFIGAGQLAYR---FTAAGILSAHKIIASSPEKMLPTVSALKRKGYNLTRSKNETVKH 58  
 DB 8 IGFIGGNMVAIGSLVRCGIVKASQVQVSP--HIENLQMRDLCAVTCDDCMVLEH 65  
 QY 59 SDVFLAVKPHIIPILDEIGADVQARHI-----VSCAAGVTISVEKKLMFOP 109  
 DB 66 SDIVFCVAPR---MLTPCAALQKVKHVPASAKDASKLVSVLADTSLTELEAFSPFGS 121  
 QY 110 AP-KVIRCMTNTPVVVGEGATVATGTHALVEDGQLLEQLMSSVGFCTEVEDLIDAVTG 168  
 DB 122 SELKVIPTMPTSMQVGECTVYTGNAVSHHDEKITHLNLGLAQVPESMIDAVTG 181  
 QY 169 LSGSPATVAFMAD--ADGVKMGLPRLAIOGAQALLGAAMKMLDSEQHPQQLKDNVC 226  
 DB 182 VAGCGPAFYVYITIEALADGVKQGVPRQMALQFAOITLGAATVLLTGKHPAVLREVC 241  
 QY 227 SPGATIHAIHFLESGFRSLINAVEASCIRREL 262  
 DB 242 SPGATIVGVHELEKGNLRSTLINAVEKSSORSABL 277

RESULT 10

ID 09HH99 PRELIMINARY: PRT; 270 AA.

AC 09HH99;

DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)

GN PROC.

OS *Methanosarcina acetivorans*.

OC Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;

OC *Methanosarcina*.

RN NCBI\_Taxid=2214;

RN 111

RP SEQUENCE FROM N.A.

RC STRAIN-CS:

RA Zhang J.K., White A.K., Kuetner H.C., Boccazzi P., Metcalf W.W.;

RT "Molecular and genetic analysis of proline biosynthesis in the  
 RT methanogenic archaeon *Methanosarcina acetivorans* C2A.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF305580; AAC22033.1; -  
 DR InterPro: IPR000304; P5CR.  
 DR Pfam: PF01089; P5CR; 1.  
 DR PROSITE: PS00521; P5CR; 1.  
 SO SEQUENCE 270 AA; 27943 MW; 50EC656AFC10B1CF CMC64;

Query Match 29.1%; Score 456; DB 1; Length 270;  
 Best Local Similarity 39.7%; Pred. No. 1.5e-28;  
 Matches 106; Conservative 52; Mismatches 97; Indels 12; Gaps 5;

QY 3 VGFIGAGQLAYR---FTAAGILSAHKIIASSPEKMLPTVSALKRKGYNLTRSKNETVKH 57  
 DB 6 IGFIGAGKMSALMOGTIKAGIVPENIGAS--DYEPFLDLQAKIGIRVSTDNNAVIR 63  
 QY 58 HSDVFLAVKPHIIPILDEIGADVQARHIIVSCAAGVTISVEKKLMFOPAPKVRICM 117  
 DB 64 ESDILFLAVKPHIIPILDEIGADVQARHIIVSCAAGVTISVEKKLMFOPAPKVRICM 120  
 QY 118 TMTPVVVGEGATVATGTHALVEDGQLLEQLMSSVGFCTEVEDLIDAVTGSGSPATA 177  
 DB 121 PNIAVTVSAASGAPKNAATPBDLKAALVFSVAGAVQVPESLMDAVTGLSGSPAFI 180  
 QY 178 FMADL--ADGVKMGLPRLAIOGAQALLGAAMKMLDSEQHPQQLKDNVCSPGATIH 235  
 DB 181 FVITAMADGAVLGRMDKSLTLAOTVLGAAMKALETGMHNGELKMDVTSAGTTIG 240

OY 236 LHFLSGGFRSLINAVEASCITRREL 262  
 DB 241 IHSLEAGICRAAFNMNAYIRASERSKEL 267

RESULT 11  
 087725 PRELIMINARY; PRT; 266 AA.

AC 087725;  
 DT 01-NOV-1998 (TREMblrel. 08, Created)  
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2).  
 GN PROC.  
 OS Clostridium sticklandii.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 OC Clostridium.  
 OC NCBI\_TaxID=1511;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HF, DSM 517;  
 RA Kenkies J., Ziehm R., Fritzsche K., Andreesen J.R., Pich A.;  
 RT "Proline biosynthesis in Clostridium sticklandii: Purification of  
 RT pyrroline-5-carboxylate reductase and characterization of the encoding  
 RT gene proc.":  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ010739; CA09332.1;  
 DR InterPro; IPR00304; PSCR.  
 DR Pfam; PF01089; PSCR; 1.  
 DR PROSITE; PS00521; PSCR; 1.  
 KW Oxidoreductase.  
 SQ SEQUENCE 266 AA; 28902 MW; 48B5AD46A5A43B65 CRC64;

Query Match 28.6%; Score 448; DB 2; Length 266;  
 Best Local Similarity 38.5%; Pred. No. 6.5e-28;  
 Matches 104; Conservative 51; Mismatches 97; Indels 18; Gaps 5;

OY 3 VGTGAGOLAYRPTAA-----GILSAHKIIASPEMNLPTVSALRKM---GVNLTNRKE 54  
 DB 5 IGTGCGNMAQAMISALVSKLIESNLIIVNSRKNII-----LEKMKNEYGITIANNIE 59  
 OY 55 TVKHSVLEFAYRPHIIPFLDEIGADVOARHIVVSCAGVYISSEKKIMAFOPAPVIR 114  
 DB 60 VAEKCDIVLEAVKONLYEMWMEKIKSVTKDKIEVSIAQKTMELFAHLGAY---AKIL 116  
 OY 115 RCMNTPVVOEGATVYATGTHALVEDGOLLLEOMSSVGFCTEVEDLIDAVTGLSGSP 174  
 DB 117 RTMPTPSWSEGMTAICPNSNIASBELLEVLKIESFGAVEIIEKLEFDVAVAVSGSSP 176  
 OY 175 AYAFMALD--ADGVKKGGLRRLAIOGAQALLGAAKMLDSQHQCOLKDWVCSFGAT 232  
 DB 177 AYVEMFLEAMADAIVIOGMRAQAYRFAQAVYGSAMVLESKSHPEGLKDWVCSFGATT 236  
 OY 233 IHALHFLSGGFRSLINAVEASCITRREL 262  
 DB 237 IEAVALVEERGRSAVIEAMKCKQSKEM 266

RESULT 12  
 09X8G1 PRELIMINARY; PRT; 284 AA.

ID 09X8G1;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE.  
 GN SCET-04C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OC NCBI\_TaxID=1902;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Seeger K.J., Harris D.;  
 RT "A set of ordered cosmid and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,  
 RA Kinashl H., Hopwood D.A.;  
 RT "A set of ordered cosmid and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL M.O. Microbiol. 21:77-96(1996).  
 DR EMBL; AL049819; CAB42663.1;  
 DR InterPro; IPR00304; PSCR.  
 DR Pfam; PF01089; PSCR; 1.  
 DR PROSITE; PS00521; PSCR; 1.  
 SQ SEQUENCE 284 AA; 29806 MW; 291D989EB4AB115 CRC64;

Query Match 27.0%; Score 423; DB 2; Length 284;  
 Best Local Similarity 40.1%; Pred. No. 7e-26;  
 Matches 108; Conservative 38; Mismatches 89; Indels 34; Gaps 5;

OY 4 GTGAGOLAYRPTAAGILSAHKIIASPEMNLPTVSALRKMGLVLR-----SNKER 55  
 DB 34 CMIGAG-----WAPADLT-----VTARRERAEDELARRHGVPTVTAEA 72  
 OY 56 VHSVLEFAYRPHIIPFLDEIGADVOARHIVVSCAGVYISSEKKIMAFOPAPVIR 115  
 DB 73 AKADPLITVPRQDMGTLDLAVHPADRLVYISGAGVTSFEEEL--ARGTPVVR 129  
 OY 116 CMNTPVVOEGATVYATGTHALVEDGOLLLEOMSSVGFCTEVEDLIDAVTGLSGSPA 175  
 DB 130 VMINTALVDENAMSVISACTHATAHLTETEIFGAVKTLRVPSQDCTALSGSPA 189  
 OY 176 YAFMALDA--DGVKKGGLRRLAIOGAQALLGAAKMLDSQHQCOLKDWVCSFGATT 233  
 DB 190 YFFYLVEAMTDAGITLGLPRDAHDLIVOSALGAAKMLRDSGEHVKLRNVTSFAGTTI 249  
 OY 234 HALHFLSGGFRSLINAVEASCITRREL 262  
 DB 250 NAIREFENHGVRAIALALEAARDRSREL 278

RESULT 13  
 09K308 PRELIMINARY; PRT; 272 AA.

ID 09K308;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2).  
 GN PROC.  
 OS Bacillus cereus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OC NCBI\_TaxID=1396;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 14579;  
 RA Ostad O.A., Agasse H., Iereclus D., Kolsto A.B.;  
 RT "20 kb DNA sequence surrounding the hblA hemolytic enterotoxin locus  
 RT from Bacillus cereus ATCC 14579.";  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 14579;  
 RX MEDLINE=20055637; PubMed=10569720;  
 RA Oktad O., Gombet M., Purnelle B., Rose M., Lereclus D., Kolsto A.B.  
 RT "Sequence analysis of three *Bacillus cereus* loci under P<sub>ICR</sub>-regulated  
 RL genes encoding degradative enzymes and enterotoxin."  
 DR Microbiology 145:3129-3138(1999).  
 DR EMBL: AJ007795; CAB69639.1; -  
 DR EMBL: AJ337785; CAB69790.1; -  
 DR InterPro: IPR000304; PSCR.  
 DR Pfam: PF01089; PSCR. 1.  
 KW Oxidoreductase.  
 SO SEQUENCE 272 AA; 29298 MW; 61006F9EA986B4C6 CRC64;

Query Match	25.5%	Score 399	DB 2	Length 272	
Best Local Similarity	34.1%	Pred. No. 5.3e-24			
Matches	93	Conservative 61	Mismatches 107	Indels 12	
				Gaps	
QY	3	VGFIGAGQLAYRFTAA	---	GLSAHKLIIASSPPMNPVY-SALRKMGNLRSKEIVK	57
Db	5	IGGFIGNCMGIAMIGMIKNIVSPDOITS	--	DLVNI LKNKNSNGYGITITINNEVAN	62
QY	58	HSDFLAVAPRPIIPILDEIGADVOAQRHIVS	SAAGVITSSYEKKLAFODAPKRVICM	117	
Db	63	SADILITIKPDLVSSYINQIKDKINDVYITIA	GKSTIKSTENE--	FERKLKVRVM	119
QY	118	TNPVVVQSGAIVYATGTHALVEGQLLEO	LMSSVGFCTEVEEDLIDAVTGLSSGPAYA	177	
Db	120	PNPFLVGGESALCFENEMYTEKIKIVLFI	NPFOFVNIENKIMADVITSSSPAY	179	
QY	178	FMALD--ADGGVKMLPRRLAIQUGAVALG	AKKMLLDSEOPQLOKNNSPGCGATIIA	235	
Db	180	YMEIEMADAALVGLMPKQAKVFAADAVL	GSKKMWLEIGHPEGLKDMVSSPGTITIA	239	
QY	236	LHFLESGFRRLLIWAVASCIIRRELQNS	ADQ	268	
Db	240	VYTLLEKGRITAIISAKKRCQKSMESST	STSK	272	

RESULT	14			
09DDX2				
AD	09DDX2	PRELIMINARY:	PRT:	248 AA.
IC	09DDX2			
DT	01-JUN-2001 (TREMblrel. 17, Created)			
DT	01-JUN-2001 (TREMblrel. 17, Last sequence update)			
DT	1110058B13JRIK (TREMblrel. 17, Last annotation update)			
DE	1110058B13JRIK. PROTEIN.			
GN	GN			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-C57BL/6J; TISSUE=EMBRYO;			
RC	MEDLINE=21085660; PubMed=1217851;			
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,			
RA	Sakoto T., Okazaki Y., Gotojiori T., Bono H., Kasunawa T., Saito R.,			
RA	Kido K., Matsuda I. A., Ashburren M., Batalov S., Casavant T.,			
RA	Fleischmann W. V., Gaasterland T., Glasi C., King B., Kochiya H.,			
RA	Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole C., Rong B., Kocchia H.,			
RA	Schirral L. M., Stuhli E., Suzuki R., Tomita M., Wagner L., Quackenbush J.			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bersh G.,			
RA	Blake J., Botfield D., Bojunga N., Carlinici P., de Bonaldo M. F.,			
RA	Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hall D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombers P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shidara Y., Storch K.-F.,			

RA Susuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.  
RA Myshakov-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsuke S.,  
RT Haseizaki Y.:  
"Functional annotation of a full-length mouse cDNA collection.",  
RL Nature 409:685-690(2001).  
DR EMBL: AK004291; BAB2352.1; -  
DR MGD: MG1:1913444; 111005B1R1K.  
DR InterPro: IPR000304; P5CR.  
DR Pfam: PFO1089; P5CR: 1.  
SQ SEQUENCE 248 AA: 26170 MW: 4AFA896396305A34 CRC64;

[illegible]

RESULT	15			
09NGS4				
ID	09NGS4	PRELIMINARY;	PRT;	254 AA.
AC	09NGS4			
DT	01-OCT-2000 (TREMBlrel. 15, Created)			
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)			
DE	PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2).			
GN	P5CR.			
OS	Leishmania donovani.			
OC	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania			
OX	NCBI_TaxID=5661;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=DD8;			
RA	Selanko L.M., Porter-Kelley J.M., Chaudhuri G.,			
RT	"Cloning and characterization of Leishmania p5CR gene."			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AF441824; AAF64050.1; "			
DR	Interpro: IPR000304; P5CR.			
DR	PIfam: PF01089; P5CR; 1.			
DR	PROSITE: PS00521; P5CR; 1.			
KW	Oxidoreductase.			
SQ	SEQUENCE 254 AA; 26760 MW; ADD1356GB89FF1B5 CRC64;			

	Query Match	24.08;	Score 376;	DB 5;	Length 254;	
	Best Local Similarity	36.68;	Pred. No. 3.3e-22;			
	Matches	96;	Conservative	49;	Mismatches	91; Indels 26; Gaps 6
Oy	18 AGILSA-----HKII---ASSPE--MLPLPVSLARKMGVNLTRSKNEFYKHSDVLFLA	65				
Db	2 AGMIAKAVLDEPETHFICNRRASTNELSLYHSSVGAV-----ELAKSDIITIG	52				
Oy	66 VKPHPIIFIDEIGADVOARHIIVSCAGAVTSSEKKALMAFGAPRVVICMTNPVVQO	125				
Db	53 VKRPGIVPLETIKEKITPSKLIVISMAAGVPATETNNC---PKPTVRVVMKNIPSPFG	109				
Oy	126 EGATVATGTTHALVEDGQLLEQLMSVGFCTEVEEDLLDAVTGSGCGPAVAFMALDA---	183				

Thu Jan 31 15:07:59 2002

us-09-912-717-1.rspt

Page 8

Db 110 EGYTSTSGNSAVTPDDEAVVMKLFCAIGKAYLWESAIHGCVGAGSSPAYVEMPEALS 169  
QY 184 DGVKMGLEPRRLAIQAGQALLGAARKMLDSEOHPCOLKDNVCSPGATIHAIHFLESOG 243  
Db 170 DGAVRGGIPRAQSYEMAAQVLTGAARKMLDSEKTPGATLKDWCSPGTTIEAVRPLEKGG 229  
QY 244 FRSLLINAVEASCIRPRELOSM 265  
Db 230 MRSSVIEAMTECMESKEFEKI 251

Search completed: January 31, 2002, 14:57:19  
Job time: 204 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

January 31, 2002, 14:53:30 ; Search time 23.41 Seconds  
(without alignments)  
993.550 Million cell updates/sec

```

Title: US-09-912-717-1
Perfect score: 1567
Sequence: 1 MSVGFIGAGQLAYRETAGI.....TPSSPKLLTRSLAIGCKND 314

```

scoring table: BLOSUM62

524463 seqs, 74073290 residues

number of calls satisfying chosen parameters: 522463

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1:	/S1D8/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2:	/S1D8/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3:	/S1D8/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4:	/S1D8/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5:	/S1D8/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6:	/S1D8/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7:	/S1D8/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8:	/S1D8/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9:	/S1D8/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10:	/S1D8/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11:	/S1D8/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12:	/S1D8/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13:	/S1D8/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14:	/S1D8/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15:	/S1D8/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16:	/S1D8/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17:	/S1D8/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18:	/S1D8/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19:	/S1D8/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20:	/S1D8/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21:	/S1D8/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22:	/S1D8/gcgdata/geneseq/geneseqp/AA2001.DAT.*

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1567	100.0	314	21	AAB20584	Human delta 1-pyrrol
2	1509	96.3	320	22	AAB74779	Human pY-CR prote
3	1506	83.3	319	22	AAB74780	Human pY-CR prote
4	1045	66.7	212	22	AAB95027	Human pY-CR prote
5	1042	66.5	255	21	AAB56540	Human protein sequ
6	537.5	34.3	274	21	AAV92517	Human prostate can
7	537.5	34.3	274	22	AAV92517	Human OXRE-14. Ho
8	534.5	34.1	274	22	AAV92517	Human protein sequ
9	512.5	32.7	276	22	AAG68956	Human protein sequ
10	487.5	31.1	256	21	AAG06978	Human dihydroxyro
11	402.5	23.7	269	22	AAG06979	Arabidopsis thalia
					AAG64106	Escherichia coli p

12	372	23.7	151	21	AAAG50980
13	370	23.6	270	22	AAAG90206
14	370	23.6	220	22	AAAB78782
15	306.5	19.6	235	22	AAAG81123
16	269	17.2	78	21	AAAG03518
17	24.5	15.4	271	22	AAAG82031
18	112.5	7.2	92	21	AAAG34996
19	104	6.6	594	14	AAAR34936
20	99.5	6.3	341	22	AAAG90789
21	99.5	6.3	341	22	AAAB79110
22	99.5	6.3	341	22	AAAB79144
23	98	6.3	381	16	AAAN01433
24	98	6.3	1551	21	AAAB10434
25	97.5	6.2	543	14	AAAR42456
26	97.5	6.2	543	17	AAAR99455
27	97.5	6.2	543	19	AAAB37053
28	97.5	6.2	543	21	AAAB10470
29	97.5	6.2	543	21	AAAB10470
30	95.5	6.1	375	21	AAAG47613
31	95.5	6.1	1455	19	AAAB86653
32	94.5	6.0	372	16	AAAM01432
33	94.5	6.0	393	21	AAAG16522
34	93.5	6.0	393	21	AAAG7616
35	93.5	6.0	542	22	AAAG64458
36	91.5	5.8	1026	18	AAW37490
37	91.5	5.8	1026	21	AAAY44757
38	89	5.7	448	22	AAAG0237
39	89	5.7	448	22	AAAB80046
40	88.5	5.6	349	21	AAAG47614
41	88.5	5.6	408	22	AAAG91550
42	88.5	5.6	722	19	AAAG55104
43	88.5	5.6	2234	21	AAAB81502
44	87.5	5.6	72	21	AAAC4897
45	87	5.6	409	21	AAAG39651
					Arabidopsis thaliana
					C. glutamicum prote
					Corynebacterium gl
					Mycobacterium tube
					Human secreted prot
					S. epidermidis opo
					Arabidopsis thaliana
					CENP-B. Homo sapi
					C. glutamicum prote
					Corynebacterium gl
					Corynebacterium gl
					Glycerol-3-phospha
					S. aggraeus galact
					Enzyme involved in
					Biosynthetic enzym
					S. putrefaciens EP
					S. putrefaciens PK
					Shewanella putrefa
					Arabidopsis thaliana
					Fanconi anemia of
					Glycerol-3-phospha
					Arabidopsis thaliana
					S. putrefaciens th
					Caulobacter cresce
					Caulobacter cresce
					C. glutamicum prote
					Corynebacterium gl
					Arabidopsis thaliana
					C. glutamicum prote
					Streptococcus pneu
					Arabidopsis thaliana
					Arabidopsis thaliana

## ALIGNMENTS

RESULT	1
ID	AAB20584
XX	AAB20584 standard; Protein; 314 AA.
XX	
AC	AAB20584;
XX	
DT	13-DEC-2000 (first entry)
XX	
DE	Human delta 1-pyrroline-5-carboxylate reductase homologue protein.
XX	
XX	
XX	Human; delta 1-pyrroline-5-carboxylate reductase homologue; P5CRH;
XX	neurotropic; neuroprotective; osteopathic; cytosolic; antiParkinsonian;
XX	cardiac; osteopathic; ophthalmological; hepatotropic; gene therapy;
XX	neuronal disorder; connective tissue disorder; cell proliferation;
XX	akathisia; Alzheimer's disease; amnesia; dementia; Parkinson's disease
XX	cytic fibrosis; myocardial fibrosis; osteoporosis; Marfan syndrome;
XX	cirrhosis; leukaemia; lymphoma; sarcoma; bone marrow cancer.
OS	Synthetic.
XX	
PN	US6100075-A.
XX	
PD	08-AUG-2000.
XX	
PF	18-JUN-1998; 98US-0099676.
XX	
PR	18-JUN-1998; 98US-0099676.
XX	
PA	(INCY-) INCYTE PHARM INC.
XX	
PI	Hillman JL, Baughn MR, Corley NC;
XX	
XX	WPI; 2000-548292/50.
DR	N-PSDB; AAA88073.
DR	

XX New human delta 1-pyrroline-5-carboxylate reductase homolog (P5CRH) and  
 PT polynucleotides encoding P5CRH, useful for diagnosing, treating or  
 PT preventing neuronal or connective tissue disorders, or disorders of  
 PT cell proliferation

PS Claim 1: Fig 1A-E; 32pp; English.

XX The present sequence represents human delta 1-pyrroline-5-carboxylate  
 CC reductase homologue (P5CRH). P5CRH has nootropic, neuroprotective,  
 CC osteopathic, cytosolic, antiparkinsonian, caldiant, osteopathic,  
 CC ophthalmological and hepatotropic activities, and can be used in gene  
 CC therapy. P5CRH and the polynucleotides encoding it are useful for  
 CC diagnosing, treating or preventing neuronal disorders, connective tissue  
 CC disorders, or disorders of cell proliferation. These may be applied to  
 CC any subject, e.g. dogs, cats, cows, horses, rabbits, monkeys, or  
 CC preferably humans. The polynucleotide sequences may be used to detect  
 CC and quantitate gene expression in biopsied tissues where expression of  
 CC the polypeptide may correlate with the disease. The diagnostic assay may  
 CC be used to determine absence, presence, and excess expression of the  
 CC polypeptide, and to monitor regulation of the polypeptide levels during  
 CC therapeutic intervention. The polypeptide or its fragments are useful  
 CC for treating or preventing neuronal disorder (e.g. akathisia,  
 CC Alzheimer's disease, amnesia, dementia or parkinson's disease),  
 CC connective tissue disorder (e.g. cystic fibrosis, myocardial fibrosis,  
 CC osteoporosis or Marfan syndrome), or a disorder of cell proliferation  
 CC (e.g. cirrhosis, leukaemia, lymphoma, sarcoma or bone marrow cancer).  
 CC The polypeptide or its fragments are also useful for screening  
 CC libraries of compounds in any of several drug-screening techniques.

XX Sequence 314 AA:

Query Match 100.0%; Score 1567; DB 21; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-159;  
 Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVGFAGAGOLAVRFTAGILSAHKTIASSPEMNLPTVSALRKMGVNLRSNKETVKHSD 60  
 DB 1 msygfagagqlayrfllaagllsahkltasspemmllpvsalrkmgvnlrsnketykhsd 60  
 QY 61 VLFIAVRPHIIPFLIDEIGADVQARHIIVSCAAGVTISSVEKKLMAFQAPKVIKMTNT 120  
 DB 61 vlfiavrphiiplfildelgadvgarhivscagvtlssvekklmalfqapkvirkmtnt 120  
 QY 121 PVVVOEGATVYATGTHALVEDGQLLEQDMSSVGFCTEVEDLIDAVTGISGSPAYAFMA 180  
 DB 121 pvvvgatvayaltghalvedgqlleqdmssvgfctevvedlidavtgisgspayafma 180  
 QY 181 LDADGVKMGIPRRLAIOIGAQLLGAAKMLDSEOHPCOLKDNVCSPGATTHALHFE 240  
 DB 181 ldadgvmkgiprrlaioigaqlllgaakmldseghpcqldkdvcspgatthalhfe 240  
 QY 241 SGGFSILINAVEASCIRTELQSMADOKETISPAALKKTLIDRVKLESPTVSTLTPSSPG 300  
 DB 241 sggsflllnaveascirtrelqsmadoketispalkkltldrvklesptvstltspspg 300  
 QY 301 KLTLSLAIGKKD 314  
 DB 301 kltlrslalgykkd 314

RESULT 2

AAB74779  
 ID AAB74779 standard; Protein; 320 AA.

XX AAB74779;

XX 06-JUN-2001 (first entry)

XX Human Py-CR protein SEQ ID NO:4.

XX Human; pyrroline 5'-carboxylate reductase; Py-CR; P5CR.

XX Homo sapiens.  
 OS  
 XX CN1274728-A.  
 PN  
 XX  
 PD 29-NOV-2000.  
 XX

PR 25-MAY-1999; 99CN-0107071.

XX 25-MAY-1999; 99CN-0107071.

XX (UYFU-) UNIV FUDAN.

XX Yu L, Fu Q, Zhang H;

XX WPI: 2001-211749/22.

XX N-PSDB; AAF81847.

XX New human protein and its code sequence, preparation and application -  
 PT Claim 4; Page 16; 20pp; Chinese.

XX The present invention describes a human protein designated Py-CR, which  
 CC is homologous to the human pyrroline-5'-carboxylate reductase (P5CR)  
 CC (EC1.5.1.2). The present invention also describes methods for the  
 CC application and production process of the Py-CR polynucleotide and  
 CC protein sequences. The present sequence represents the human Py-CR  
 CC protein as given in the present invention.

XX Sequence 320 AA:

Query Match 96.3%; Score 1509; DB 22; Length 320;  
 Best Local Similarity 96.9%; Pred. No. 3.3e-153;  
 Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2;

QY 1 MSVGFAGAGOLAVRFTAGILSAHKTIASSPEMNLPTVSALRKMGVNLRSNKETVKHSD 56  
 DB 1 msygfagagqlayrfllaagllsahkltasspemmllpvsalrkmgvnlrsnketykhsd 60  
 QY 57 KHSVDLEAVKPHIIPFLIDEIGADVQARHIIVSCAAGVTISSVEKKLMAFQAPKVIKIRC 116  
 DB 61 khsvdleavkphiiplfildelgadvgarhivscagvtlssvekklmalfqapkvirk 120  
 QY 117 MTNPPVVOEGATVYATGTHALVEDGQLLEQDMSSVGFCTEVEDLIDAVTGISGSPAY 176  
 DB 121 mtnppvvgatvayaltghalvedgqlleqdmssvgfctevvedlidavtgisgspay 180  
 QY 177 AFMADLD--ADGVKMGIPRRLAIOIGAQLLGAAKMLDSEOHPCOLKDNVCSPGATTH 234  
 DB 181 afmadldadgvmkgiprrlaioigaqlllgaakmldseghpcqldkdvcspgatth 240  
 QY 235 ALHFIESGFSILINAVEASCIRTELQSMADOKETISPAALKKTLIDRVKLESPTVSTL 294  
 DB 241 alhfiesgfslllnaveascirtrelqsmadoketispalkkltldrvklesptvstl 300  
 QY 295 TPSSPGKLTLSLAIGKKD 314  
 DB 301 tpsspgkltlrslalgykkd 320

RESULT 3

AAB74780  
 ID AAB74780 standard; Protein; 319 AA.

XX AAB74780;

XX 06-JUN-2001 (first entry)

XX Human P5CR protein.

XX Human; pyrroline 5'-carboxylate reductase; Py-CR; P5CR.

OS Homo sapiens.  
 XX CNI274728-A.  
 XX 29-NOV-2000.  
 XX 25-MAY-1999: 99CN-0107071.  
 XX 25-MAY-1999: 99CN-0107071.  
 XX (UYFU-) UNIV FUDAN.  
 XX Yu L, Fu Q, Zhang H;  
 XX MPI: 2001-211749/22.  
 XX  
 XX New human protein and its code sequence, preparation and application -  
 XX Example 2; Fig 1: 20pp; Chinese.  
 XX  
 XX The present invention describes a human protein designated Py-CR, which  
 XX is homologous to the human pyrroline-5'-carboxylate reductase (P5CR)  
 XX (ECL 5.1.2). The present invention also describes methods for the  
 XX application and production process of the Py-CR polynucleotide and  
 XX protein sequences. The present sequence represents the human P5CR  
 XX protein which is used in comparison with the human Py-CR protein in  
 XX an example from the present invention.  
 XX  
 XX Sequence 319 AA:  
 SQ  
 Query Match 83.3%; Score 1306; DB 22; Length 319;  
 Best Local Similarity 82.7%; Pred. No. 1.9e-131;  
 Matches 263; Conservative 27; Mismatches 22; Indels 6; Gaps 2;  
 QY 1 MSVGTIGAGOLAYR---FTAGILSAHKIIASPEENMLPTVSALRRKMGVNLTRSKETV 56  
 Db 1 msvgtigagolafafalagfytaagvtaahkimasppmdlatvaltkmgvkltpnhketv 60  
 QY 57 KHSVDLFLAVKPHIIPILDEIGADVQARIIVYSCAAGVITSSVEKKLMAFOFARVINC 116  
 Db 61 qhsdvlflavkphliffldelgadedrhlvscagavtlsslekklafppaprvlvc 120  
 QY 117 MTNTPVVOEGATYATGTHALVEDGQLLEQLMSSVGFCTEVEDBIDAVTGSGSGPAY 176  
 Db 121 mtntpvvvregatvayalgthaqvedgrlmeqllstvtgctveeedlidaetvlgsgspay 180  
 QY 177 AFNALD--ADGVYKMGPRRLAIOQLGQALIGAAKMLDSEQHPCOLKDNVCSFGCATIH 234  
 Db 181 afnaldldagvymgprlrlaioqlgqaligaaakmlhseqhpgqlkdnvsgpactih 240  
 QY 235 ALHFLSEGGFRSLINAVEASCIPTRELQSMADQEKISPAALKKTLTLDVKLESTVSTL 294  
 Db 241 alhvlseggfrslilnaveascirrelqsmadqevspaalkktildvklidspactal 300  
 QY 295 TPSSPGKLTLSLALGK 312  
 Db 301 spsgkltlprslapagk 318  
 RESULT 4  
 ID AAB93027 standard; Protein: 212 AA.  
 XX AAB93027;  
 XX  
 XX 26-JUN-2001 (first entry)  
 XX  
 XX Human protein sequence SEQ ID NO:11795.  
 XX  
 XX Human; primer: detection; diagnosis; antisense therapy; gene therapy.  
 XX  
 XX Homo sapiens.

XX  
 XX EPI074617-A2.  
 XX  
 XX 07-FEB-2001.  
 XX  
 XX 28-JUL-2000: 2000EP-0116126.  
 XX  
 XX 29-JUL-1999: 99JP-0248036.  
 XX 27-AUG-1999: 99JP-0300253.  
 XX 11-JAN-2000: 2000JP-0118776.  
 XX 02-MAY-2000: 2000JP-0183767.  
 XX 09-JUN-2000: 2000JP-0241899.  
 XX  
 XX (HELI-) HELIX RES INST.  
 XX  
 XX Ota T, Isega T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 XX Ishii S, Sugiyama T, Wakematsu A, Nagai K, Otsuki T;  
 XX MPI: 2001-318749/34.  
 XX  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 XX full-length cDNAs defined in the specification, and for the detection  
 XX and/or diagnosis of the abnormality of the proteins encoded by the  
 XX full-length cDNAs -  
 XX  
 XX Claim 8; SEQ ID 11795; 2537pp + CD ROM; English.  
 XX  
 XX The present invention describes primer sets for synthesizing 5602  
 XX full-length cDNAs defined in the specification. Where a primer set  
 XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 XX to the complementary strand of a polynucleotide which comprises one of  
 XX the 5602 nucleotide sequences defined in the specification, where the  
 XX of an oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 XX complementary strand of a polynucleotide which comprises a 5'-end  
 XX sequence and an oligonucleotide comprising a sequence complementary to a  
 XX polynucleotide which comprises a 3'-end sequence, where the  
 XX the 5'-end sequence/3'-end sequence is selected from those defined in  
 XX the specification. The primer sets can be used in antisense therapy and  
 XX particularly full-length cDNAs. The primers are also useful for the  
 XX detection and/or diagnosis of the abnormality of the proteins encoded by  
 XX the full-length cDNAs. The primers allow obtaining of the full-length  
 XX cDNAs easily without any specialized methods. AAH03166 to AAH13628 and  
 XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 XX represent oligonucleotides, all of which are used in the exemplification  
 XX of the present invention.  
 XX  
 XX Sequence 212 AA:  
 SQ  
 Query Match 66.7%; Score 1045; DB 22; Length 212;  
 Best Local Similarity 99.1%; Pred. No. 1e-103;  
 Matches 210; Conservative 0; Mismatches 0; Indels 2; Gaps 1;  
 QY 105 MAFQAPKVIKMTNTPVVOEGATYATGTHALVEDGQLLEQLMSSVGFCTEVEDBID 164  
 Db 1 mafqapkvirkmtntpvvvegatvayalgthaqvedgqlleqlmssvgctveeedlid 60  
 QY 165 AVTGLSGSGPAYAFMALD--ADGVYKMGPRRLAIOQLGQALIGAAKMLDSEQHPCOLK 222  
 Db 61 avtglssgspayafmaldaadagvymgprlrlaioqlgqaligaaakmlidseqhpgqlk 120  
 QY 223 DNVCSFGCATIHAFLESGFRSLINAVEASCIPTRELQSMADQEKISPAALKKTLTLD 282  
 Db 121 dnvcsfgcatihaflesggfrslilnaveascirrelqsmadqekispaalkktilld 180  
 QY 283 RYKLESPVSTLTSSPGKLTLSLALGKKD 314  
 Db 181 rylvlespvtstltcsspgkltltslalgkkd 212







KW Human; dihydropyrirole-5-carboxylate reductase 30; cancer; cytostatic;  
 KW human immunodeficiency virus; HIV; infection; immunological disease;  
 KW inflammatory disease.

OS Homo sapiens.

XX CNI298002-A.

XX 06-JUN-2001.

XX 24-NOV-1999; 99CN-0124090.

XX 24-NOV-1999; 99CN-0124090.

XX (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.

XX Mao Y, Xie Y;

XX MPI: 2001-489680/54.

XX N-PSDB; AAH77597.

XX Human dihydropyrirole-5-carboxylate reductase 30 as one new kind of  
 PT polypeptide and polynucleotides encoding this polypeptide -  
 PS Claim 1; Page 20-21 (disclosure); 26pp; Chinese.

CC The invention relates to a novel polypeptide, human  
 CC dihydropyrirole-5-carboxylate reductase 30, polynucleotides encoding  
 CC this polypeptide and a DNA recombination process to produce the  
 CC polypeptide. The polypeptide is useful for treating various diseases,  
 CC such as malignant tumours, nosohaemia, HIV infection, immunological  
 CC diseases and inflammatory diseases. The invention also provides an  
 CC antibody against the polypeptide. The present sequence is the  
 CC polypeptide of the invention.

XX Sequence 274 AA;

SQ

Query Match 34.1%; Score 534.5; DB 22; Length 274;  
 Best Local Similarity 44.9%; Pred. No. 8.2e-49;  
 Matches 120; Conservative 46; Mismatches 88; Indels 13; Gaps 5;

QY 3 VGEFGAGG---LAVRTAGILSAHKIIASSP-EMNLPTVSALRKGVNLTRSNKRETVK 57  
 Db 11 VGIYVAGRMAGATAGGIIAGKVEAQHILASAPDNLCHFGAL---GRTIHSNQEVILQ 67  
 QY 58 HSDVFLAVKPHIIPFILDIDIGADVQARHTVSCAAGVTISVVEKKLMAFOPAPKVIKCM 117  
 Db 68 scIIVfIetkphvIpaVlaevapvttehlIIVaaqmsIstleelI---pntIrvIrvI 124  
 QY 118 TMTPTVVOEGATVYATGTHALVDEGQLLEQLMSSVGCCTEVEBDLIDAVTGLSGSPAY- 176  
 Db 125 pIIPcvvqgaIvmargrhvyssetkIILqHleacgrceevpeayvdihtgIsgsvatV 184  
 QY 177 -AFMLDDDGKMGKLPRLAIQLGAQALLGAAKMLDSEQHPCQLKDNVCSFGATIFA 235  
 Db 185 cefsaaleagvkmqmpslahriaqlllgctakmlIhegnhpqqlsvdctpggtIiyg 244  
 QY 236 LHFLESGGFRSLILNAVASCIRTEL 262  
 Db 245 lhaIegggIraatmsaveatcraekel 271

RESULT 9

AAAG06978

XX AAG06978 standard; Protein; 276 AA.

XX AAG06978;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 3954.

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126284.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 04-MAY-1999; 99US-0132048.  
 PR 05-MAY-1999; 99US-0132407.  
 PR 06-MAY-1999; 99US-0132485.  
 PR 07-MAY-1999; 99US-0132486.  
 PR 11-MAY-1999; 99US-0132487.  
 PR 14-MAY-1999; 99US-0132863.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.





```

PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 14-OCT-1999; 99US-0158328.
PR 14-OCT-1999; 99US-0158330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 31.1%; Score 487.5; DB 21; Length 256;  
 Best Local Similarity 41.6%; Pred. No. 8.1e-44;  
 Matches 107; Conservative 51; Mismatches 92; Indels 7; Gaps 3;

```

QY 8 AGOLAVFTAGILSAHKTIASSPENMLPTVSALRRKGVNLTNRNKEVYKHSDFLAVK 67
DB 2 Aeslaagvaasgvlpnr-ictavhenlrrdyfaefgvnvtseeevkedvifsvk 60
QY 68 PHIIPLDEIGADVOARHIVSCAGVITSSVEKKLAFQAPRVITRCMTNPVVOEG 127
DB 61 pyvvkaveelckeklnkllivsvaagikindlqe---wsgqdfitvmpnptaaevga 116
QY 128 ATYVAAGTAAHVEDGQLLEQLMSVSGFCTEVEDLIDAVTGISGSGPAYARMAID--ADG 185
DB 117 asvmslgtateedgavlmfegavgkllkadekmfdavtgisgsgpaylilaaleadg 176
QY 186 GVMGPRRLAIOLOLGAQALGAAKMLLSEOHPCOLKDNVCSPGGATIHALLPLESGFR 245

```

```

DB 177 gvaaglprelalslaeqvlgaaetmvsctgkhpvgkdvstpggtciagvhelekstr 236
QY 246 SLINAVEASCIRREL 262
DB 237 atlmavvaakstrel 253

```

## RESULT 11

AAG64106 standard; protein; 269 AA.

AC AAG64106;

DT 21-SEP-2001 (first entry)

DE Escherichia coli proc polypeptide.

KM L-pipecolic acid production; delta-1-piperidine-6-carboxylic acid;  
 KM pyrrolidine-5-carboxylate reductase; lysine-6-aminotransferase;  
 KW proc.

OS Escherichia coli.

PN WO200148216-A1.

PD 05-JUL-2001.

PF 22-DEC-2000; 2000WO-JP09137.

PR 28-DEC-1999; 99JP-0373389.

(SAOC ) MERCIAN CORP.

PI Fujii T, Aritoku Y, Mukaihana M, Narita T, Agematu H, Isehiki K;

DR WPT: 2001-418286/44.

DR N-PSDB; AAH48471.

PT Producing L-pipecolic acid for use as a pharmaceutical raw material,  
 PT comprises reducing delta-1-piperidine-6-carboxylic acid with  
 PT pyrrolidine-5-carboxylate reductase particularly in recombinant  
 PT bacterium

PS Disclosure; Page 31-33; 54pp; Japanese.

CC The invention relates to a method for producing L-pipecolic acid.

CC The method comprises reducing delta-1-piperidine-6-carboxylic acid with  
 CC pyrrolidine-5-carboxylate reductase. A recombinant Escherichia coli or  
 CC coryneform bacterium that can express a gene encoding

CC lysine-6-aminotransferase is cultured in an L-lysine-containing  
 CC medium in order to produce L-pipecolic acid. L-pipecolic acid may be  
 CC used as a pharmaceutical raw material. The process is efficient, and  
 CC since it is conducted enzymatically, chirality of the final product  
 CC can be retained. The present sequence is the Escherichia coli proc  
 CC polypeptide. The gene encoding this protein may be used in the  
 CC invention.

XX Sequence 269 AA;

Query Match 25.7%; Score 402.5; DB 22; Length 269;  
 Best Local Similarity 35.5%; Pred. No. 1.1e-34;  
 Matches 98; Conservative 52; Mismatches 107; Indels 19; Gaps 6;

```

QY 3 VGTIGQALVFTAGILSAHKTIASSPENMLPT---VSAIR-KMGVNLTRSNKEVYKH 58
DB 5 lfigcgmgg-kallgylasgvlpigklywtpdpkvaahdqfglnaesqevaqi 63
QY 59 SDVLFVAVPHIIPPLDEIGADVOARHIVSCAGVITSSVEKKLAFQAPRVITRCMT 118
DB 64 adlifaavkpgymikvliseltsslnkdsilvsvlaagvldqlar---alghdrKliramp 120

```





DB	116	---	vvrvmpmptclmlyvgkmgstvtlkgryvdagleykdkllstvgdvleavesdldavtam	172
QY	110	AKRVLRKMTNTFVVVQEGTVATGTHALVEGQLLEQMLSSVGFTETEEDPLDAVTVL	169	
DB	56	sgaadeadvflvcfkpkflveitgtgldnnsaqsvvmaasqslstaameesasqip	115	
QY	53	KETVKSVDVFLAVNPHIIPITLIDETGADV---QARHIYSCAAGYTSISVEKKLMAFP	109	
DB	3	tlavvgggdgig-ealvsgliaan----mmpq-nirvtntseergelrdtygllmt-dn	55	
QY	2	SVGFAGACOLARFRTAAGILSAHKILIASSPENMLPTVSALRRMG-----VNLTRSN	52	
DB	3	tlavvgggdgig-ealvsgliaan----mmpq-nirvtntseergelrdtygllmt-dn	55	
QY	23.6%	Score 370;	DB 22;	Length 270;
	Best Local Similarity	33.8%;	Pred. No. 3.5e-31;	
	Matches	93;	Conservative	56;
			Mismatches	102;
			Indels	24;
			Gaps	
XX	XX	AAE71753 to AAE72330 encode the Corynebacterium glutamicum metabolic		
XX	XX	pathway (WP) proteins given in AAE79634 to AAB80211. The C. glutamicum		
XX	XX	MP nucleic acids are useful for the production of fine chemicals		
XX	XX	in microorganisms, including organic acids, nonproteinogenic amino		
XX	XX	acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,		
XX	XX	saturated and unsaturated fatty acids, diols, carbohydrates, aromatic		
XX	XX	compounds, vitamins, cofactors, polyketides and enzymes.		
XX	XX	Sequence	270 AA;	
XX	XX	Claim 20; Page 589-590; 1737pp; English.		
XX	XX	Nucleic acids from Corynebacterium glutamicum encoding metabolic		
XX	XX	pathway proteins, useful for producing fine chemicals in		
XX	XX	microorganisms, including organic acids, nonproteinogenic amino acids,		
XX	XX	and purine and pyrimidine bases -		
XX	XX	WPI: 2001-137957/14.		
XX	XX	N-PSDB; AAE71901.		
XX	XX	(BADI ) BASF AG.		
XX	XX	Pompejus M, Kroegeer B, Schroeder H, Zelder O, Habenerauer G;		
XX	XX	PR 09-JUL-1999; 99DE-1032186.		
XX	XX	PR 09-JUL-1999; 99DE-1032206.		
XX	XX	PR 09-JUL-1999; 99DE-1032227.		
XX	XX	PR 09-JUL-1999; 99DE-1032228.		
XX	XX	PR 09-JUL-1999; 99DE-1032229.		
XX	XX	PR 09-JUL-1999; 99DE-1032230.		
XX	XX	PR 14-JUL-1999; 99DE-1032922.		
XX	XX	PR 14-JUL-1999; 99DE-1032926.		
XX	XX	PR 14-JUL-1999; 99DE-1032928.		
XX	XX	PR 14-JUL-1999; 99DE-1033004.		
XX	XX	PR 14-JUL-1999; 99DE-1033005.		
XX	XX	PR 14-JUL-1999; 99DE-1033006.		
XX	XX	PR 12-AUG-1999; 99DE-1048613.		
XX	XX	PR 27-AUG-1999; 99DE-1040764.		
XX	XX	PR 27-AUG-1999; 99DE-1040765.		
XX	XX	PR 27-AUG-1999; 99DE-1040766.		
XX	XX	PR 27-AUG-1999; 99DE-1040832.		
XX	XX	PR 31-AUG-1999; 99DE-1041378.		
XX	XX	PR 31-AUG-1999; 99DE-1041379.		
XX	XX	PR 31-AUG-1999; 99DE-1041380.		
XX	XX	PR 31-AUG-1999; 99DE-1041394.		
XX	XX	PR 31-AUG-1999; 99DE-1041396.		
XX	XX	PR 03-SEP-1999; 99DE-1042077.		
XX	XX	PR 03-SEP-1999; 99DE-1042079.		
XX	XX	PR 03-SEP-1999; 99DE-1042086.		
XX	XX	PR 03-SEP-1999; 99DE-1042087.		
XX	XX	PR 03-SEP-1999; 99DE-1042088.		
XX	XX	PR 03-SEP-1999; 99DE-1042095.		
XX	XX	PR 03-SEP-1999; 99DE-1042124.		
XX	XX	PR 03-SEP-1999; 99DE-1042129.		
XX	XX	PR 09-MAR-2000; 2000US-0187970.		



OY 170 SCGSPAYAFMALDA--DGVKMGLEPRRLAIOGAOALLGAAKMLDSEOHPCOLKDNVCS 227  
 DB 173 sgsspaylflytcealltaagynlgypratlakklavaafegaaatmmketykpsellragvas 232  
 OY 228 PGGATTIHALHLESGFRSLINAVEASCTIRREL 262  
 DB 233 pagtvealreleesgltgafyraqacdrseel 267

RESULT 15  
 AAG81123  
 ID AAG81123 standard; Protein; 295 AA.  
 XX  
 AC AAG81123:  
 XX  
 DN 04-SEP-2001 (first entry)  
 XX  
 DE Mycobacterium tuberculosis potential drug target protein SEQ ID 174.  
 XX  
 KM Drug target; growth; organism viability; characterisation.  
 XX  
 OS Mycobacterium tuberculosis.  
 XX  
 PN WO200135317-A1.  
 XX  
 PD 17-MAY-2001.  
 XX  
 PF 13-NOV-2000; 2000WO-US31152.  
 XX  
 PR 12-NOV-1999; 99US-0165086.  
 PR 12-NOV-1999; 99US-0165124.  
 PR 01-FEB-2000; 2000US-0179531.  
 XX  
 PA (RECC ) UNITV CALIFORNIA.  
 XX  
 PI Eisenberg D, Rotstein SH, Marcotte EM;  
 XX  
 DR WPI: 2001-329193/34.  
 DR N-PSDB; AAH51974.  
 XX  
 XX

Identifying nucleotide or polypeptide sequence for use as drug target,  
 involves providing algorithm that analyzes a functional relationship  
 between nucleotide or polypeptide sequences, and comparing the  
 sequences -

Disclosure; Page 161; 207pp; English.

This invention relates to a method for identifying a nucleotide or  
 polypeptide sequence that may be a drug target, or essential for growth  
 or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092  
 represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium  
 tuberculosis proteins which are potential drug targets. The DNA and  
 protein sequences are used to illustrate the method of the invention. The  
 method involves providing an unknown nucleotide or polypeptide sequence,  
 and comparing it to a number of sequences along with at least one  
 algorithm capable of analysing a functional relationship between  
 nucleotide and polypeptide sequences. The method is useful for  
 characterising the function of nucleic acids and polypeptides that may be  
 useful as a target for a drug or essential for the growth or viability of  
 an organism.

Sequence 295 AA;

Query Match 19 68; Score 306.5; DB 22; Length 295;  
 Best Local Similarity 34.3%; Pred. No. 2.6e-24;  
 Matches 86; Conservative 39; Mismatches 91; Indels 35; Gaps 6;

OY 48 LTRSNKEFYKHSVDVLEAVKP-HIIPFLDEGADV-----QAKHIYSCAAGVTISSVE 101  
 DB 54 lvsaaadaventalftvavkpadvprvialanaadaaandsaeqvtvvaqiltlayfe 113

OY 102 KILMAFQAPAKVIRCTNTPTVVYQEGATVYATGTHALVEDGOLLEQIMSSVGECTEVEED 161  
 DB 114 sklpagcp---vvvampnaaalvgaytalakgrfvtppqldeevsaifdaygvjltvpes 170  
 OY 162 LIDAVTGLSGSPAYAFMALDA--DGVKMGLEPRRLAIOGAOALLGAAKMLDSEOH-- 217  
 DB 171 qldavtavsgspayffllvaelvdagvvgjlsrgvaldlaaqtmagsaamltermedq 230  
 OY 218 -----PCOLKDNVCSPGAT IHALHLESGFRSLINAVEASCTIRREL 262  
 DB 231 ggangelmglrvdltaarilaavtspgcltaaalrelergfrtmaavdaavga----- 283  
 OY 263 QSMADQEKISP 273  
 DB 284 ksrseqllrltp 294

Search completed: January 31, 2002, 14:54:44  
 Job time: 74 sec

Thu Jan 31 15:07:53 2002

us-09-912-717-1.rag

---